DOCUMENT RESUME

ED 056 090 24 TM 000 896

AUTHOR Green, Donald Ross
TITLE Racial and Ethnic Bias in Test Construction. Final Report.
INSTITUTION CTB/McGraw Hill, Monterey, Calif.
SPONS AGENCY Office of Education (DHEW), Washington, D.C.
BUREAU NO BR-0-I-033
PUB DATE 24 Sep 71
CONTRACT OEC-9-70-0058(057)
NOTE 104p.

EDRS PRICE MF-$0.65 HC-$6.58
DESCRIPTORS Achievement Tests; Caucasians; Data Analysis; Disadvantaged Youth; Ethnic Groups; Grade 1; Grade 3; Grade 5; Grade 8; Grade 10; *Item Analysis; *Mexican Americans; Minority Group Children; *Negroes; Racial Discrimination; Rural Youth; Suburban Youth; *Test Bias; *Test Construction; Test Reliability; Test Validity; Urban Youth
IDENTIFIERS California Achievement Tests

ABSTRACT

To determine if tryout samples typically used for item selection contribute to test bias against minority groups, item analyses were made of the California Achievement Tests using seven subgroups of the standardization sample: Northern White Suburban, Northern Black Urban, Southern White Suburban, Southern Black Rural, Southern White Rural, Southwestern Mexican Urban and Southwestern Anglo-American Suburban. The best half of the items in each test were selected for each group. Typically about 30% of the items in the upper half of the distribution of item-test correlations for a group on a test did not meet this criterion with another group. By this criterion minority groups were relatively similar as were the three suburban groups. The resulting unique item tests did not correlate well with each other. Scores of minority groups were relatively better on the selected items. Thus, standard item selection procedures produce tests best suited to groups like the majority of the tryout sample and are therefore biased against other groups to some degree. This degree varies. Ways to minimize this bias need to be developed. (Author/MS)

ED056090

PA 24
BR 0-I-033
TM

Final Report

Project No. 0-I-033
Contract No. OEC-9-70-0058 (057)

Donald Ross Green
CTB/McGraw-Hill
Del Monte Research Park
Monterey, California 93940

RACIAL AND ETHNIC BIAS IN TEST CONSTRUCTION

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

September 24, 1971

TM 000 896

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
OFFICE OF EDUCATION
REGIONAL RESEARCH PROGRAM

ABSTRACT

Problem: Do the tryout samples typically used for item selection contribute to test bias against minority groups?

Method: Item analyses were made of the *California Achievement Tests* using seven subgroups of the standardization sample: Northern White Suburban, Northern Black Urban, Southern White Suburban, Southern Black Rural, Southern White Rural, Southwestern Mexican Urban and Southwestern Anglo-American Suburban. The best half of the items in each test were selected for each group.

Results: Typically about 30% of the items in the upper half of the distribution of item-test correlations for a group on a test did not meet this criterion with another group. By this criterion minority groups were relatively similar as were the three suburban groups. The resulting unique item tests did not correlate well with each other. Scores of minority groups were relatively better on the selected items.

Conclusions: Standard item selection procedures produce tests best suited to groups like the majority of the tryout sample and are therefore biased against other groups to some degree. This degree varies. Ways to minimize this bias need to be developed.

Final Report

Project No. 0-I-033
Contract No. OEC-9-70-0058 (057)

Racial and Ethnic Bias in Test Construction

Donald Ross Green

CTB/McGraw-Hill

Monterey, California

September 24, 1971

The research reported herein was performed pursuant to a contract with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Regional Research Program

TABLE OF CONTENTS

	Page
Abstract	ii
Acknowledgements	v
List of Tables	vi
Problem and Objectives	1
Problem	1
Related Literature	2
Objectives of the Study	4
Limitations	4
Method	4
Sample	5
Data Analysis	7
Results	8
Proportion of Unique Items	8
Independence of Unique Item Tests	12
Changes in Test Characteristics	12
Changes in Test Scores	21
Conclusions	21
Recommendations	23
Bibliography	24
Appendix	27

ACKNOWLEDGMENTS

Special thanks are due to my colleague, Dr. George Burket, who first suggested this use of item selection procedures and who offered valuable advice throughout. I would also like to thank Dr. Joseph L. Dionne who gave encouragement and who ensured that CTB/McGraw-Hill provided the facilities and personnel needed at a cost substantially in excess of that provided by the grant. Finally, I would like to thank Lois Maclin who kept track of the entire project, made a number of statistical analyses, constructed all the tables, and edited the entire report.

LIST OF TABLES

		Page
1.	Characteristics of the sample groups.	6
2.	Proportion of items selected which are unique for the four pairs of groups by grade.	9
3.	Median proportion of unique items by grade and test.	11
4.	Median porportion of unique items for all possible pairs of groups.	11
5.	Correlations between scores on unique item tests.	13
6.	Estimated minimum proportion of unique item test variances which are independent.	14
7.	Frequencies of changes in reliability, whole test to half-test, showing greater gain (or less loss) than shown for the comparison group.	16
8.	Median item-test correlations for four tests.	17
9.	Median item-test correlations of items selected for the original whole test and for the resulting half-test.	18
10.	Frequencies of increases and decreases in point biserial correlations for items selected for four tests.	20
11.	Frequencies of changes, whole test to half-test, in mean differences showing relatively higher scores for the comparison group.	22
12.	Frequencies of increases and decreases in relative differences between unique item test means: majority test to minority test.	22
13.	Proportion of items selected which are unique for all possible comparison groups for Grade 1.6.	27
14.	Proportion of items selected which are unique for all possible comparison groups for Grade 3.6.	28
15.	Proportion of items selected which are unique for all possible comparison groups for Grade 5.6.	29
16.	Proportion of items selected which are unique for all possible comparison groups for Grade 8.6.	30
17.	Proportion of items selected which are unique for all possible comparison groups for Grade 10.6.	31
18.	Reliabilities and relationships between items unique to Groups I & II, Grade 1.6.	32
19.	Reliabilities and relationships between items unique to Groups I & II, Grade 3.6.	33
20.	Reliabilities and relationships between items unique to Groups I & II, Grade 5.6.	34
21.	Reliabilities and relationships between items unique to Groups I & II, Grade 8.6.	35
22.	Reliabilities and relationships between items unique to Groups III & IV, Grade 1.6.	36
23.	Reliabilities and relationships between items unique to Groups III & IV, Grade 3.6.	37
24.	Reliabilities and relationships between items unique to Groups III & IV, Grade 5.6.	38

List of Tables (continued)

		Page
25.	Reliabilities and relationships between items unique to Groups III & IV, Grade 8.6.	39
26.	Reliabilities and relationships between items unique to Groups III & IV, Grade 10.6.	40
27.	Reliabilities and relationships between items unique to Groups IV & V, Grade 1.6.	41
28.	Reliabilities and relationships between items unique to Groups IV & V, Grade 3.6.	42
29.	Reliabilities and relationships between items unique to Groups IV & V, Grade 5.6.	43
30.	Reliabilities and relationships between items unique to Groups IV & V, Grade 8.6.	44
31.	Reliabilities and relationships between items unique to Groups IV & V, Grade 10.6.	45
32.	Reliabilities and relationships between items unique to Groups VI & VII, Grade 1.6.	46
33.	Reliabilities and relationships between items unique to Groups VI & VII, Grade 3.6.	47
34.	Reliabilities and relationships between items unique to Groups VI & VII, Grade 5.6.	48
35.	Reliabilities and relationships between items unique to Groups VI & VII, Grade 8.6.	49
36.	Means and crossover means on unique item tests for Groups I & II.	50
37.	Means and crossover means on unique item tests for Groups III & IV.	52
38.	Means and crossover means on unique item tests for Groups IV & V.	55
39.	Means and crossover means on unique item tests for Groups VI & VII.	58
40.	Means and standard deviations on the whole and half-tests for northern white suburban versus northern black urban, Grade 1.6.	60
41.	Means and standard deviations on the whole and half-tests for northern white suburban versus northern black urban, Grade 3.6.	61
42.	Means and standard deviations on the whole and half-tests for northern white suburban versus northern black urban, Grade 5.6.	62
43.	Means and standard deviations on the whole and half-tests for northern white suburban versus northern black urban, Grade 8.6.	63
44.	Means and standard deviations on the whole and half-tests for southern white suburban versus southern black urban, Grade 1.6.	64
45.	Means and standard deviations on the whole and half-tests for southern white suburban versus southern black rural, Grade 3.6.	65
46.	Means and standard deviations on the whole and half-tests for southern white suburban versus southern black rural, Grade 5.6.	66
47.	Means and standard deviations on the whole and half-tests for southern white suburban versus southern black rural, Grade 8.6.	67

List of Tables (continued)

		Page
48.	Means and standard deviations on the whole and half-tests for southern white suburban versus southern black rural, Grade 10.6.	68
49.	Means and standard deviations on the whole and half-tests for southern black rural versus southern white rural, Grade 1.6.	69
50.	Means and standard deviations on the whole and half-tests for southern black rural versus southern white rural, Grade 3.6.	70
51.	Means and standard deviations on the whole and half-tests for southern black rural versus southern white rural, Grade 5.6.	71
52.	Means and standard deviations on the whole and half-tests for southern black rural versus southern white rural, Grade 8.6.	72
53.	Means and standard deviations on the whole and half-tests for southern black rural versus southern white rural, Grade 10.6.	73
54.	Means and standard deviations on the whole and half-tests for southwestern Mexican-American versus southwestern Anglo-American, Grade 1.6.	74
55.	Means and standard deviations on the whole and half-tests for southwestern Mexican-American versus southwestern Anglo-American, Grade 3.6.	75
56.	Means and standard deviations on the whole and half-tests for southwestern Mexican-American versus southwestern Anglo-American, Grade 5.6.	76
57.	Means and standard deviations on the whole and half-tests for southwestern Mexican-American versus southwestern Anglo-American, Grade 8.6.	77
58.	Reliability coefficients and standardized reliability coefficients on the whole and half-tests for Groups I & II, Grade 1.6.	78
59.	Reliability coefficients and standardized reliability coefficients on the whole and half-tests for Groups I & II, Grade 3.6.	79
60.	Reliability coefficients and standardized reliability coefficients on the whole and half-tests for Groups I & II, Grade 5.6.	80
61.	Reliability coefficients and standardized reliability coefficients on the whole and half-tests for Groups I & II, Grade 8.6.	81
62.	Reliability coefficients and standardized reliability coefficients on the whole and half-tests for Groups III & IV, Grade 1.6.	82
63.	Reliability coefficients and standardized reliability coefficients on the whole and half-tests for Groups III & IV, Grade 3.6.	83
64.	Reliability coefficients and standardized reliability coefficients on the whole and half-tests for Groups III & IV, Grade 5.6.	84
65.	Reliability coefficients and standardized reliability coefficients on the whole and half-tests for Groups III & IV, Grade 8.6.	85

List of Tables (continued)

Page

66. Reliability coefficients and standardized reliability coefficients on the whole and half-tests for Groups III & IV, Grade 10.6. 86
67. Reliability coefficients and standardized reliability coefficients on the whole and half-tests for Groups IV & V, Grade 1.6. 87
68. Reliability coefficients and standardized reliability coefficients on the whole and half-tests for Groups IV & V, Grade 3.6. 88
69. Reliability coefficients and standardized reliability coefficients on the whole and half-tests for Groups IV & V, Grade 5.6. 89
70. Reliability coefficients and standardized reliability coefficients on the whole and half-tests for Groups IV & V, Grade 8.6. 90
71. Reliability coefficients and standardized reliability coefficients on the whole and half-tests for Groups IV & V, Grade 10.6. 91
72. Reliability coefficients and standardized reliability coefficients on the whole and half-tests for Groups VI & VII, Grade 1.6. 92
73. Reliability coefficients and standardized reliability coefficients on the whole and half-tests for Groups VI & VII, Grade 3.6. 93
74. Reliability coefficients and standardized reliability coefficients on the whole and half-tests for Groups VI & VII, Grade 5.6. 94
75. Reliability coefficients and standardized reliability coefficients on the whole and half-tests for Groups VI & VII, Grade 8.6. 95

PROBLEM AND OBJECTIVES

Problem

The standardized achievement and intelligence tests used in schools are often said to be biased against, and thus inappropriate for, children belonging to disadvantaged racial and ethnic minorities. If this is so there are two possible sources of such bias. The first may come from the preconceptions and thought patterns of the test item writers. The second may come from the customary item selection procedures used in test construction. This second possible source of bias is the general topic investigated in this study.

The typical procedure in building standardized achievement and aptitude tests--essentially unchanged over many years (cf. Lord & Novick 1968, Chapter 15; Ruch 1929, Chapter 2)-- is first to develop a pool of items meeting various specifications as to form and content. Next these items are given to a sample of individuals--the step in question here. Various item statistics, such as point biserial correlations (item vs. total score), are calculated and the "best" items are then chosen, with "best" being characterized first and foremost by a high relationship of the item to the total score. Other characteristics such as difficulty and the effectiveness of distractors (in multiple choice tests) are also considered. Most of these latter item characteristics are related to the item-test correlation to some degree. Therefore the items which "discriminate" best, i.e., show the highest relationship to total score, are the ones usually chosen. This in turn means that the characteristics or attributes of the individuals in the tryout sample which are most responsible for differences in total score determine which items tend to be chosen and determine, in effect, what it is the test measures within the range of possibilities available in the item pool. That is, certain qualities, attitudes, knowledge or skills found in varying degrees in the tryout sample will have the largest differential effect on total score on the tryout test. The items most sensitive to these attributes of the tryout sample then get selected.

Consequently, the possibility exists that the items selected are biased and discriminate against groups not adequately represented in the tryout samples. If there are traits of some atypical groups not found in the tryout sample which interact more strongly with the items than do the attributes they share with the majority, or if the group is uniformly low on these latter common traits, but not on other equally relevant attributes, then in either case, one could say the resulting test is biased. In the first instance it is biased because it measures different things for different groups unbeknownst to the users; in the second instance it measures only a portion of the relevant behaviors but is taken to measure them all.

If this is all true, then the use of "average" item tryout samples may result in the selection of item sets unsuited to one or more of the

various racial, ethnic, cultural minority groups in our schools. It may be noted in passing that if all this is true it may also follow that no single tryout group can ever solve the problem--perhaps only the construction of separate tests would do so although this has obvious drawbacks. Another alternative might be to use the same test but different item weights for different groups.

This study attempts to estimate how important the matter of tryout samples is. Specifically, would the use of samples drawn from minority groups for item tryouts result in the selection of different items? It is customarily assumed that the choice of subjects for item tryouts is not very important, although "atypical" groups (such as disadvantaged children) are usually avoided. Some evidence for evaluating this assumption is presented in this report.

Related Literature

Prior work in this area does not seem to have dealt directly with this particular issue. In fact, as far as achievement tests are concerned, very little work of any sort on the matter of bias appears to be available. The work on intelligence and aptitude is more extensive, but other aspects of the bias issue than the one considered here have dominated discussions. The present study concerns achievement tests, but since the problem is essentially the same--as are the tests in many ways (Kelley 1927)--the intelligence test studies are relevant.

That children's intelligence test scores are related to their social and economic status was reported by Binet and others almost 60 years ago and has been studied and argued about ever since. For a long time these arguments largely stayed within the bounds of the much older and highly emotional nature-nurture controversy, perhaps because many felt that the then new tests could settle that argument (Terman 1916, pp. 19-20). Since the intensity of those arguments shows no sign of diminishing after 50 years (Jensen 1969), that hope may be considered unreasonable. In any case, the score differences favoring the more privileged elements of society remain a fact (Coleman et al. 1966). It may be added that the accusations of the misuse and the misinterpretation of scores (Hunter & Rogers 1967) are also factual in some, if not most, instances.

However, the issue here is the nature of the tests themselves. This has not been as widely studied as it might be. Apparently, the first serious attempt to examine test items for bias was led by Allison Davis and his colleagues 20 years ago (Eells et al. 1951). They examined several existing group intelligence tests and the items in them in an attempt to determine the factors built into the tests related to differences in performance between cultural groups. They concluded: "Variations in opportunity for familiar cultural words, objects, or processes required for answering the test items seem . . . the most adequate

general explanation. . .(Eells 1951, p. 68)." This sort of objection to standard tests continues to be made (Wasserman 1969).

Interestingly, the subjects in the Eells study were all white and drawn from the schools of "a western industrial city of about 100,000 people." One result of the study was the publication of the Davis-Eells Games (1953) which was designed to eliminate this kind of cultural bias. Three things may be noted about this test. First, the test--now out of print--proved to yield as substantial differences between SES groups (Angelino & Shedd 1955) as other group intelligence tests. Second, they eliminated the items that showed SES differences in difficulty provided they could rationalize the difference as a consequence of opportunity. Lastly, they apparently did not look at the differences between SES groups with respect to item discrimination. The common interpretation of the outcome of the Davis-Eells test and similar efforts by others has been that the task of building a "culture free" or "culture fair" test may be not only impossible but inappropriate because the test so made would not be valid, as indeed was the case for the Davis-Eells Games (e.g., Lorge 1966).

Anastasi (1968) points out that while this conclusion is proper there is still the issue of bias in prediction, and in recent years the assertions that group intelligence tests discriminate against various minority and disadvantaged groups in our society have increased in number and vehemence. Some school systems (New York City, for example) have virtually abandoned the use of such tests (Gilbert 1966). Similarly some college personnel now argue that the various placement and ability tests traditionally used are inappropriate (Brown & Russell 1964). Many of these arguments are sound, but those that claim the tests fail to function among disadvantaged minority students in the way they do in other groups lack supporting evidence. A series of studies at both the high school and college levels show that academic aptitude tests predict grades just as well in such groups as they do among more privileged groups (e.g., Stanley & Porter 1967; Temp 1971). Only the work of Green and Farquhar (1965) points to a different conclusion among a half dozen or so studies on this issue. In fact some work even points the other way, i.e., it suggests that some scholastic aptitude tests over-predict the performance of lower class and Negro students in contrast to middle class and white students (Hewer 1965; Cleary 1968). Data obtained by Kennedy et al. (1963) show that the grandfather of them all, the Stanford-Binet (Terman & Merrill 1960), produces equal or higher item-test correlations for an all black southern sample than was reported in either the 1937 or the 1960 standardization.

Still, there are racial and cultural differences in ability patterns as shown so clearly by Lesser, Fifer, & Clark (1965), and hence the possibility continues to exist that tests based on items selected for a particular group (such as black, ghetto children) would be less biased and more useful for them. Bias in the sense of faulty prediction is now beginning to be studied extensively (Cleary & Hilton 1968; Linn & Werts 1971) but bias in tests not designed to predict has not yet been really explored.

Objectives of the Study

A number of problems occur when trying to consider bias in achievement tests because the criteria of bias are not crystal clear. Most recent writers (Cardall & Coffman 1964; Cleary & Hilton 1968; Potthoff 1966; Messick and Anderson 1970; Green 1971) say something about tests which measure different things when used with different groups.

Two ways in which this could occur as a consequence of the customary tryout samples and item selection procedures were noted in the general statement of the problem above. This suggests that the characteristics of tryout samples may be more important than usually acknowledged. To explore this possibility this study compares the results of using three disadvantaged minority groups--northern urban black, southern rural black, western Mexican-American--as tryout samples in contrast to white advantaged groups in the same regions.

The study attempts to determine (a) if these different groups would lead to the selection of different items from the item pool; if so, (b) do the different items selected measure different things; and (c) are the resulting item sets selected "better" for the minority groups in the sense that they are more reliable and have better functioning items (higher point biserial correlations); (d) if the relative discrepancy in scores favoring majority groups would be reduced by using a minority tryout group.

Limitations

The major limitation of this study lies in the restricted nature of the item pool; all items come from an already published test. They are therefore preselected and may be limited in their possibility of eliciting differential reactions from the sample groups. A pool of items written with this purpose in mind would have been better. Another limitation is the somewhat uncontrolled nature of the samples. The set from which the schools were chosen was randomly selected but the specific schools used were those appearing to meet certain criteria most closely; in this sense the selection was arbitrary. Third it should be noted that grade and test level are not independent; the test levels were designed to be continuous and articulate well but they are different tests. Thus the assumption made throughout the material below that grade differences are meaningful may not be justified. Finally, because of limitations of time and money not all relevant analyses of the data could be made.

METHOD

The basic data for this study were derived from that obtained during the standardization of the *California Achievement Tests, 1970*

Edition, (CTB McGraw-Hill 1970). The *California Achievement Tests* (CAT) have as their purpose the measurement of educational attainment and the provision for analysis of learning difficulties. They are basically similar to the 1957 edition and generally measure:

(1) the ability to understand the meaning of the content material presented,

(2) the performance of the student in applying rules, facts, concepts, conventions, and principles to solve problems in the basic curricular material, and

(3) the level of performance of the student in using the tools of reading, mathematics, and language in progressively more complicated situations.

CAT is a general achievement test battery with five overlapping levels. The tests in the battery which were investigated in this study are Reading Vocabulary, Reading Comprehension, Total Reading, Mathematics Computation, Mathematics Concepts and Problems, Total Mathematics, Language Mechanics, Language Usage and Structure, and Total Language. Total Reading, Total Mathematics, and Total Language were treated as tests separate from their parts. The standardization took place early in 1970 and involved over 200,000 students in about 400 schools. The sampling design called for obtaining a sample of school districts stratified by region (seven areas), school district size (three categories by average enrollment per grade), community type (urban, town, rural rated by density), and control (public or parochial). Within the districts, schools were chosen randomly for each test level, and all students in the selected schools who were in appropriate grades took the test.

The items in the test came from a variety of sources but it is fair to say that they were written by and for "middle America." The tryout samples also fit this description. Thus the test should favor white middle-class Americans if it favors any group.

Sample

All schools participating in this standardization of CAT answered questionnaires which provided information on the basic character of the area served (e.g., residential suburb, inner part of a large city, etc.), the percentage of white students, the percentage of children from homes where another language is spoken, and the percentage of children in families falling in each of four SES groups defined by parental occupation (professional-managerial, white collar, skilled worker, unskilled).

From data on these questionnaires seven groups of schools were drawn for this study from which four pairs were made for comparisons. The groups characteristics and sizes are shown in Table 1. Thus the samples used in this study are drawn from schools serving pupils highly

Table 1

CHARACTERISTICS OF THE SAMPLE GROUPS

Group Number	Group Label	Geographic[a] Region	Residential Type	Ethnic Group	Socioeconomic Status	Number of Cases by Grade				
						Grade 1.6	3.6	5.6	8.6	10.6
I	NWS	North	Residential Suburban	White (97%)[b]	High (81%)[b]	299	225	265	328	--
II	NBU	North	Central City	Black (99%)	Low (81%)	285	304	278	250	--
III	SWS	South	Residential Suburban	White (99%)	High (77%)	361	211	293	304	279
IV	SBR	South	Rural	Black (100%)	Low (96%)	202	220	171	245	183
V	SWR	South	Rural	White (91%)	Low (81%)	323	200	199	296	246
VI	WMU	Southwest	Small and Large Cities	Mexican-American[c] (87%)	Low (82%)	146	144	169	399	--
VII	WAS	Southwest	City and Suburban	Anglo-American (99%)	High (81%)	189	218	249	277	--

[a] The states containing these particular school systems are: 1) North: Illinois, Kansas, New Jersey, Indiana; 2) South: Georgia, South Carolina, Alabama; 3) Southwest: Arizona, Texas, Oklahoma.

[b] Estimated per cent of cases falling in the category.

[c] 81% speak mostly Spanish at home.

homogeneous with respect to ethnic background and rather homogeneous with respect to socioeconomic status. Only at grade 10 was it not possible to always find schools meeting these criteria in the standardization population; sufficiently segregated tenth grades were found only in the South.

The groups were paired for comparisons as follows:

a. Northern, black, central city (NBU) versus Northern white suburban (NWS).
b. Southern black rural (SBR) versus Southern white suburban (SWS).
c. Southern black rural (SBR) versus Southern white rural (SWR).
d. Southwestern Mexican-American (WMU) versus Southwestern Anglo-American suburban (WAS).

Enough schools meeting the appropriate criteria to provide between 150 and 300 students for each group at each of five grade levels were selected. The grade levels are 1, 3, 5, 8 and 10 so that each of the five levels of the CAT battery is included. Grade 10 comparisons were made in the South only as noted above. Most comparisons were made for each of the nine tests of the battery listed above.

Data Analysis

Four sorts of analyses were made.

(1) The basic procedure used for examining the data was an item selection routine. Each of the seven groups was treated as a tryout sample with the items in each test functioning as an item pool. The "best" half of the items for each group were identified by noting those items with the highest item-test correlations within the group covered. The number and per cent of items chosen for one member of a pair of groups but not the other was recorded. The number of these "unique" items indicate the degree to which different groups interact in a unique manner with the test items. All 21 possible pairs of groups were compared this way; all the remaining analyses were made only for the four pairs listed above.

(2) Scores for each group in a pair were obtained on both sets of unique items and the correlation between the two sets of scores was found; from these correlations, estimates of the variance not common to the two unique item "tests" were made to judge how different the sets of items really are in what they measure. Thus this analysis supplements the first.

(3) Another analysis consisted of examining and comparing whole test and half-test KR 20 reliability estimates since differential reliability would be a form of bias indicating that the test scores have a larger error component in one group than they do in another group.

(4) Finally, mean scores on whole, half, and unique tests were examined for changes in relative status of the groups as a result of item selection.

The identification of the items with the highest point biserial correlations was done separately for each of the seven groups on each of the nine kinds of tests, each having five levels. Note that the Total Reading, Mathematics, and Language tests were treated separately from their parts. Actually, there were only 274 separate analyses rather than the 315 possible ones (7 x 9 x 5) because of the missing Grade 10 data noted above and because no analyses were made of the Total Language scores in grades 1, 5, and 8 for the NWS group, and in grades 1 and 8 in the NBU group.

RESULTS

Proportion of Unique Items

The results of these analyses were a series of tests "best" for each group, half as long as the original test--when N was odd the expression (N+1)/2 was used to determine the length of the half-test. The next step was to identify those items selected for only one of the two members of a pair. Obviously, the number of such items is the same for both groups. This number as a proportion of the items in these half-tests is an index of the degree to which the item selection procedure produces a different test for the two groups. Table 2 exhibits these proportions for the four basic comparison groups.

The overall median proportion is approximately .30. The proportions do not appear to vary systematically by grade or test (see Table 3). However, certain groups appear considerably more like each other than are others by the criterion of the relative size of these proportions. It can be readily seen from Table 2 that the WMU-WAS (VI vs. VII) groups differ more than do the other 3 pairs. Groups SBR (IV) and SWR (V) differ the least. These proportions of unique items for all 21 comparisons can be found in Tables 13-17 in the Appendix.

The medians of these proportions for the various possible pairs are shown in Table 4. As expected, the white middle-class groups are consistently more like each other (these pairs have lower medians) than they are like the minority groups. The latter also have more in common than they share with the three majority groups. The SWR group (V) does not fully fit into this otherwise clear pattern. However, in general they appear more like the three minority groups than they resemble the three suburban groups. Of course, economically they are undoubtedly more disadvantaged than these three, albeit much less so than the southern black group.

Table 2

PROPORTION OF ITEMS SELECTED WHICH ARE UNIQUE
FOR THE FOUR PAIRS OF GROUPS BY GRADE

Grade 1.6	Number of Items Selected	Groups Compared I & II	III & IV	IV & V	VI & VII
Vocabulary	46	41	33	35	59
Comprehension	12	25	58	33	42
Reading Total	58	40	36	34	69
Computation	20	15	25	40	25
Concepts & Problems	24	42	38	42	58
Math Total	44	16	25	23	41
Mechanics	19	42	21	21	58
Usage & Structure	10	30	30	40	40
Language Total	37	--	24	27	54
Grade 3.6					
Vocabulary	20	30	65	35	45
Comprehension	23	22	26	22	35
Reading Total	43	28	42	28	33
Computation	36	17	28	22	25
Concepts & Problems	23	35	48	35	43
Math Total	59	29	32	30	32
Mechanics	33	48	42	30	45
Usage & Structure	13	31	46	23	46
Language Total	46	41	30	28	48
Grade 5.6					
Vocabulary	20	50	55	35	70
Comprehension	21	48	43	29	52
Reading Total	41	46	46	37	61
Computation	34	41	38	21	41
Concepts & Problems	20	50	40	20	55
Math Total	54	44	46	20	46
Mechanics	40	45	35	25	53
Usage & Structure	21	33	48	38	33
Language Total	61	--	30	16	26

Table 2 Continued

Grade 8.6	Number of Items Selected	Groups Compared I & II	III & IV	IV & V	VI & VII
Vocabulary	20	40	15	15	45
Comprehension	23	22	39	30	39
Reading Total	43	26	23	21	44
Computation	24	25	46	29	29
Concepts & Problems	25	36	40	36	28
Math Total	49	29	49	35	29
Mechanics	36	42	33	42	39
Usage & Structure	25	36	56	32	16
Language Total	61	--	15	15	18
Grade 10.6					
Vocabulary	20	--	55	40	--
Comprehension	23	--	22	22	--
Reading Total	43	--	42	30	--
Computation	24	--	33	33	--
Concepts & Problems	25	--	40	32	--
Math Total	49	--	33	24	--
Mechanics	40	--	38	35	--
Usage & Structure	27	--	41	30	--
Language Total	67	--	21	19	--
Median for all tests and all grades		35	38	30	43

Table 3

MEDIAN PROPORTION OF UNIQUE ITEMS BY GRADE AND TEST

Test	Grade 1.6	3.6	5.6	8.6	10.6	All Grades
Vocabulary	33	40	45	25	40	35
Comprehension	42	26	33	30	22	30
Reading Total	36	30	41	26	30	33
Computation	25	25	29	29	33	25
Concepts & Problems	38	35	35	28	32	35
Math Total	30	29	37	29	24	29
Mechanics	32	33	35	35	35	33
Usage & Structure	40	31	33	28	33	33
Language Total	30	30	21	18	19	24
ALL TESTS BY GRADE	33	30	35	29	30	30

Table 4

MEDIAN PROPORTION OF UNIQUE ITEMS FOR ALL POSSIBLE PAIRS OF GROUPS

Group	GROUP I	II	III	IV	V	VI	VII
I	--	36	26	35	30	38	26
II	36	--	33	26	25	25	41
III	26	33	--	38	30	33	27
IV	35	26	38	--	30	30	41
V	30	25	30	30	--	24	33
VI	38	25	33	30	24	--	43
VII	26	41	27	41	33	43	--

Independence of the Unique Item Tests

All groups differ from each other and some of the differences appear to be substantial. However, it is possible that these sets of "unique" items still measure pretty much the same thing. To check out this possibility, scores for each individual were obtained on both sets of unique items. This was possible since each individual answered all items. The correlations between these two scores were obtained for each group on each test (see Table 5). The number of unique items was very small in many cases. Consequently, the reliabilities are low. Full data on these unique item tests can be found in Tables 18 through 53 in the Appendix.

One way to avoid exaggerating the apparent lack of relationship between the measures because of low reliability is to correct for attenuation; the result is a figure (r'_{xy}) which is an index of the maximum amount of relationship possible given completely reliable tests. But we are interested in the degree to which scores on the pairs of unique tests vary independently, that is, in their lack of relationship.

The square of the correlation (r^2_{xy}) is an estimate of the common variance and the difference between that figure and one ($1-r^2_{xy}$) is then an estimate of the proportion of the variance of these scores that occurs independently. Thus, ($1-r'^2_{xy}$) is an estimate of the minimum proportion of independent variation in the two sets of scores. Table 6 shows such estimates for the unique item tests. Since these are minimum estimates, it clearly follows that in many instances they are measuring quite different things and as a rule do so for both groups involved.

Changes in Test Characteristics

The data examined so far are concerned with the possibility of bias stemming from items which measure different things when used with different groups. A special case of this kind of bias occurs if the test scores of one group contained substantially more error than they do for another group. The reliabilities for each group by test can be found in Tables 36 through 53 in the Appendix. The overall median

Table 5

CORRELATIONS BETWEEN SCORES ON UNIQUE ITEM TESTS

Tests	Grades 1.6		3.6		5.6		8.6		10.6	
	I	II	I	II	I	II	I	II		
Vocabulary	.34	.37	.40	.46	.49	.25	.57	.54		
Comprehension	.25	.03	.69	.55	.55	.36	.43	.27		
Reading Total	.42	.38	.79	.64	.60	.34	.51	.43		
Computation	.49	.38	.45	.50	.35	.39	.51	.44		
Concepts & Problems	.48	.53	.64	.49	.35	.39	.51	.44		
Math Total	.41	.40	.72	.63	.47	.51	.66	.54		
Mechanics	.46	.41	.62	.58	.50	.48	.47	.58		
Usage & Structure	.42	.15	.41	.33	.12	.20	.19	.16		
Language Total			.65	.53						
	III	IV	III	IV	III	IV	III	IV	III	IV
Vocabulary	.54	.77	.35	.57	.65	.51	.49	.33	.63	.35
Comprehension	.09	.72	.42	.61	.60	.40	.59	.29	.40	.19
Reading Total	.50	.78	.57	.76	.72	.57	.65	.46	.70	.34
Computation	.57	.77	.39	.43	.43	.42	.51	.20	.58	.66
Concepts & Problems	.54	.76	.42	.76	.39	.48	.62	.49	.58	.36
Math Total	.62	.79	.60	.68	.60	.59	.68	.50	.64	.59
Mechanics	.33	.56	.67	.59	.53	.62	.64	.66	.62	.68
Usage & Structure	.27	.60	.34	.33	.32	.17	.18	.08	.35	.07
Language Total	.56	.79	.54	.48	.58	.48	.46	.43	.39	.40
	IV	V	IV	V	IV	V	IV	V	IV	V
Vocabulary	.67	.63	.54	.62	.48	.61	.33	.44	.37	.47
Comprehension	.50	.30	.61	.55	.41	.52	.30	.35	.02	.28
Reading Total	.67	.65	.71	.74	.65	.73	.47	.55	.44	.56
Computation	.73	.67	.63	.68	.41	.42	.23	.54	.69	.67
Concepts & Problems	.77	.68	.71	.67	.37	.46	.44	.57	.27	.44
Math Total	.76	.69	.82	.81	.58	.58	.42	.65	.56	.58
Mechanics	.69	.24	.53	.59	.57	.62	.58	.60	.58	.54
Usage & Structure	.57	.39	.25	.49	.08	-.02	-.04	.14	.09	.29
Language Total	.69	.42	.57	.70	.42	.61	.37	.40	.43	.43
	VI	VII	VI	VII	VI	VII	VI	VII		
Vocabulary	.34	.46	.53	.47	.44	.66	.42	.44		
Comprehension	.10	.61	.61	.64	.32	.53	.49	.61		
Reading Total	.20	.54	.53	.65	.51	.73	.58	.62		
Computation	.32	.41	.50	.52	.37	.45	.39	.45		
Concepts & Problems	.47	.68	.64	.44	.33	.34	.39	.40		
Math Total	.58	.67	.52	.61	.35	.45	.50	.62		
Mechanics	-.17	.64	.54	.63	.46	.59	.63	.61		
Usage & Structure	.19	.50	.22	.37	-.08	.17	.06	.20		
Language Total	.15	.60	.43	.60	.21	.48	.33	.39		

Table 6

ESTIMATED MINIMUM PROPORTION OF UNIQUE ITEM TEST VARIANCES WHICH ARE INDEPENDENT

Tests	Grades 1.6		3.6		5.6		8.6		10.6	
	I	II	I	II	I	II	I	II		
Vocabulary	.79	.77	.76	.55	.46	.84	.07	.20		
Comprehension	--	1.00	--	.15	.27	.44	--	.56		
Reading Total	.67	.75	.02	.26	.38	.74	.35	.54		
Computation	.28	.68	.31	.54	.72	.70	.22	.45		
Concepts & Problems	.39	--	--	.31	.69	.49	.29	.18		
Math Total	.55	.49	.06	.36	.60	.53	.12	.30		
Mechanics	.62	.61	.48	.39	.62	.55	.58	.47		
Usage & Structure	--	.43	.44	.55	--	.07	.33	.85		
Language Total			.42	.45						
	III	IV	III	IV	III	IV	III	IV	III	IV
Vocabulary	.37	.14	.78	.44	--	.38	.17	--	.13	.47
Comprehension	.93	.13	.41	--	--	.54	.17	.31	.29	--
Reading Total	.42	.19	.21	.06	.05	.38	.06	.05	.10	.51
Computation	.20	--	.71	.73	.64	.57	.48	.88	.14	.20
Concepts & Problems	.07	--	--	.03	.63	.44	.17	.06	.05	.64
Math Total	--	.06	.14	.34	.40	.45	.45	.46	.29	.43
Mechanics	.67	.21	.26	.48	.49	.04	.30	.24	.30	.34
Usage & Structure	--	--	.55	.53	--	--	--	.85	--	.94
Language Total	--	--	.47	.60	--	.20	.33	.33	.71	.62
	IV	V	IV	V	IV	V	IV	V	IV	V
Vocabulary	.40	.36	.31	.28	.13	.06	.15	.17	--	.28
Comprehension	.45	.45	--	.20	.23	--	.21	.60	.92	.38
Reading Total	.43	.32	.10	.03	.09	.04	--	.16	--	.15
Computation	.27	.37	.22	.06	.52	.59	.65	.35	.14	.08
Concepts & Problems	--	.05	.06	.04	.39	.17	.14	--	.73	.28
Math Total	.08	.14	--	--	.32	.38	.49	.27	.39	.30
Mechanics	.16	.74	.49	.42	.28	.29	.51	.42	.52	.55
Usage & Structure	.29	.50	--	--	.89	.98	.97	.61	.75	--
Language Total	.27	.44	.39	.15	.23	--	.51	.44	.54	.66
	VI	VII	VI	VII	VI	VII	VI	VII		
Vocabulary	.64	.65	.41	.54	.62	.15	.59	.60		
Comprehension	--	.06	.09	.07	.66	.35	.18	.06		
Reading Total	.90	.62	.49	.17	.52	.11	.40	.37		
Computation	.80	.67	.46	.30	.69	.57	.70	.57		
Concepts & Problems	.18	.09	.03	.22	.70	.60	.46	.27		
Math Total	.28	.25	.58	.30	.80	.62	.51	.11		
Mechanics	.91	.45	.33	.35	.55	.47	.32	.37		
Usage & Structure	.86	--	.86	.39	.93	.67	.96	--		
Language Total	.92	.42	.66	.45	.87	.48	.72	.66		

KR 20's for groups I through VII are .91, .91, .91, .92, .93, .90, and .92 respectively. Obviously, there is little evidence of bias by this criterion, although a test-by-test comparison of these reliabilities shows that the figures are mostly higher for the majority group (97 of 162 comparisons).

Another way of considering this matter is to examine the changes in reliability stemming from selecting the "best" half of the items. Since the items chosen for these half-tests were those with the highest point biserial correlations for that group, the shorter tests should not be as relatively unreliable as one might otherwise expect. Because reliability is a function of the number of items, all the reliabilities on both whole and half-tests were converted or standardized to what they would become on 100 item tests (100-item-KR 20 = $100\,r/[N + (100-N)r]$ where "r" is the original reliability and N is the original number of items).

These 100 item KR 20's are shown in Tables 58 through 75 in the Appendix. With few exceptions, they are high and adequate. However, the point of interest is the relative change in these figures for the comparison groups when the "best" items for the respective groups are selected. If the original set contains much bias in the sense that it lacks reliability, one might expect that the half-test KR 20's would increase more (or decrease less) relative to the figure for the full test for the group against which it is biased. Assessing these 100 item KR 20's on this basis produces the data in Table 7. In general, the result favors a hypothesis of bias against the minority groups, but this effect does not appear until Grade 5 for the white versus black comparisons. In Grades 1 and 3 the reverse tends to be true for these pairs. While these changes in relative size do not appear to be attributable to chance, they are large in only a few instances and cannot therefore be treated as very important.

A related question is what happens to the individual items when they are part of a "best set." In particular, are the point biserial correlations (item-test correlations) better for the items when part of the half-test than when part of the whole test? When the *California Achievement Test* was originally constructed, the minimum accepted point biserial for items included was 0.25; naturally, a number of items performed less well for the groups in this study, and many of these item test correlations (about 10%) were less than 0.25. The median correlations (and/or cutting points) for four of the tests are shown in Table 8.

The median point biserials for the best half of the items selected from these tests are shown in Table 9 along with the medians for these same items when related to half-test scores. As Table 9 suggests, these correlations in the half-tests do tend to improve modestly, about .02, indicating that the resulting half-tests are slightly better for the various groups after selection. Furthermore, this improvement occurs for most items in all groups at all grade levels in each of the four tests checked, as can be seen in Table 10. However, not only is the

Table 7

FREQUENCIES OF CHANGES IN RELIABILITY, WHOLE TEST TO HALF-TEST,[a]
SHOWING GREATER GAIN (OR LESS LOSS) THAN SHOWN FOR THE COMPARISON GROUP

	Groups Compared: Majority vs. Minority																	
Grade	I vs. II				III vs. IV				V vs. IV				VII vs. VI				All Pairs				χ^2	p
1.6	5	3			8	1			8	1			2	7			23	12			3.5	NS
3.6	5	4			8	1			5	4			3	6			21	15			1.0	NS
5.6	1	7			0	9			4	5			0	9			5	30			17.8	.001
8.6	2	6			0	9			1	8			3	6			6	29			15.1	.001
10.8	--	--			0	9			1	8			--	--			1	17			16.0	.001
All Grades	13	20			29	16			26	19			8	28			56	103			13.9	.001
χ^2	1.5				3.8				1.1				11.1				13.9					
p	NS				NS				NS				.001				.001					

[a]All reliabilities were standardized to 100-item length before differences were

[b]Equal changes were classified as favoring the majority group

Table 8

MEDIAN ITEM-TEST CORRELATIONS
FOR FOUR TESTS

Group	I	II	III	IV	V	VI	VII
Grade 1.6							
Vocabulary	.323	.353	.394	.576	.434	.301	.395
Comprehension	.389	.323	.367	.593	.443	.197	.556
Computation	.549	.558	.520	.647	.614	.524	.572
Mechanics	.476	.526	.422	.653	.414	.247	.566
Grade 3.6							
Vocabulary	.594	.540	.452	.497	.537	.508	.524
Comprehension	.579	.474	.473	.522	.522	.463	.534
Computation	.451	.511	.370	.505	.483	.423	.433
Mechanics	.509	.420	.455	.476	.471	.416	.468
Grade 5.6							
Vocabulary	.448	.410	.463	.423	.475	.416	.453
Comprehension	.406	.385	.449	.401	.414	.352	.409
Computation	.353	.402	.394	.411	.442	.391	.342
Mechanics	.417	.375	.413	.419	.475	.336	.403
Grade 8.6							
Vocabulary	.388	.448	.501	.411	.490	.445	.445
Comprehension	.387	.334	.436	.304	.403	.371	.443
Computation	.450	.438	.459	.375	.481	.485	.458
Mechanics	.362	.444	.513	.477	.450	.449	.444
Grade 10.6							
Vocabulary	--	--	.438	.333	.376	--	--
Comprehension	--	--	.392	.277	.350	--	--
Computation	--	--	.440	.533	.489	--	--
Mechanics	--	--	.416	.508	.459	--	--

Table 9

MEDIAN ITEM-TEST CORRELATIONS OF
ITEMS SELECTED FOR THE ORIGINAL WHOLE TEST
AND FOR THE RESULTING HALF-TEST

	Vocabulary		Comprehension		Computation		Mechanics	
	Whole	Half	Whole	Half	Whole	Half	Whole	Half
Group I								
Grade 1.6	.414	.450	.488	.508	.674	.685	.582	.649
Grade 3.6	.674	.733	.674	.688	.580	.617	.594	.584
Grade 5.6	.549	.535	.467	.499	.472	.468	.494	.479
Grade 8.6	.499	.481	.456	.490	.519	.549	.444	.429
Group II								
Grade 1.6	.435	.446	.408	.474	.624	.642	.580	.645
Grade 3.6	.584	.619	.592	.624	.617	.604	.510	.534
Grade 5.6	.469	.524	.440	.463	.519	.508	.499	.488
Grade 8.6	.524	.538	.417	.449	.549	.538	.543	.529
Group III								
Grade 1.6	.429	.467	.424	.474	.621	.699	.599	.674
Grade 3.6	.519	.533	.570	.592	.508	.563	.594	.624
Grade 5.6	.499	.479	.524	.508	.474	.463	.466	.483
Grade 8.6	.574	.581	.481	.513	.549	.567	.558	.558
Grade 10.6	.517	.549	.433	.458	.521	.549	.486	.506
Group IV								
Grade 1.6	.555	.578	.605	.674	.654	.724	.706	.735
Grade 3.6	.542	.563	.555	.583	.588	.580	.574	.563
Grade 5.6	.494	.517	.490	.547	.483	.508	.521	.505
Grade 8.6	.499	.513	.399	.385	.404	.467	.574	.556
Grade 10.6	.383	.449	.449	.474	.620	.645	.622	.617
Group V								
Grade 1.6	.549	.538	.524	.574	.658	.688	.463	.481
Grade 3.6	.584	.608	.620	.604	.567	.574	.545	.565
Grade 5.6	.538	.569	.458	.524	.574	.599	.538	.569
Grade 8.6	.570	.599	.430	.499	.549	.534	.481	.510
Grade 10.6	.435	.499	.411	.433	.563	.590	.539	.530

Table 9 (Continued)

MEDIAN ITEM-TEST CORRELATIONS OF
ITEMS SELECTED FOR THE ORIGINAL WHOLE TEST
AND FOR THE RESULTING HALF-TEST

	Vocabulary		Comprehension		Comprehension		Mechanics	
	Whole	Half	Whole	Half	Whole	Half	Whole	Half
Group VI								
Grade 1.6	.386	.413	.288	.274	.583	.613	.308	.338
Grade 3.6	.540	.569	.511	.535	.563	.599	.513	.505
Grade 5.6	.496	.508	.467	.492	.495	.513	.436	.457
Grade 8.6	.542	.563	.449	.474	.511	.542	.492	.478
Group VII								
Grade 1.6	.481	.517	.613	.674	.688	.774	.699	.774
Grade 3.6	.608	.749	.599	.611	.539	.556	.543	.567
Grade 5.6	.513	.506	.470	.454	.421	.413	.461	.485
Grade 8.6	.479	.549	.591	.517	.504	.538	.514	.540

Table 10

FREQUENCY OF INCREASES AND DECREASES IN POINT BISERIAL CORRELATIONS FOR ITEMS SELECTED FOR FOUR TESTS

	Group I		II		III		IV		V		VI		VII	
Grade 1.6	I	D[a]	I	D	I	D	I	D	I	D	I	D	I	D
Vocabulary	37	9	34	11	35	11	35	11	37	8	34	12	32	14
Comprehension	11	1	10	2	11	1	11	1	11	1	6	5	11	1
Computation	17	2	17	3	13	7	12	8	18	2	15	5	16	4
Mechanics	14	5	17	2	15	4	14	5	15	4	11	8	15	4
Grade 3.6														
Vocabulary	17	3	19	1	16	4	18	2	20	0	18	2	17	3
Comprehension	21	2	18	3	21	2	20	3	21	2	20	3	19	3
Computation	24	12	29	7	24	11	24	11	22	12	28	7	30	6
Mechanics	20	13	25	7	21	11	23	10	19	13	25	7	17	15
Grade 5.6														
Vocabulary	15	5	18	2	18	1	13	7	16	4	17	3	16	4
Comprehension	17	3	18	3	16	4	18	3	15	5	16	5	17	4
Computation	24	10	21	13	27	6	22	11	25	9	28	6	26	8
Mechanics	23	17	29	10	33	7	26	14	27	12	30	9	28	12
Grade 8.6														
Vocabulary	17	3	20	0	20	0	18	2	18	2	20	-	18	2
Comprehension	20	3	20	3	22	1	18	5	19	4	21	1	22	-
Computation	16	8	20	1	18	6	19	4	21	3	17	7	19	5
Mechanics	25	11	24	12	24	12	27	9	29	7	32	4	25	10
Grade 10.6														
Vocabulary					16	4	16	4	20	0				
Comprehension					22	1	21	2	20	3				
Computation					20	4	19	5	22	1				
Mechanics					27	13	31	9	28	11				

[a] I = Increase
D = Decrease

size of the increase unimpressive, but its uniformity across the groups prevents one from inferring the presence of substantial bias.

Changes in Test Scores

A third way to look at bias is to assert that the scores of some groups are unfairly low because the test does not adequately measure all the relevant abilities or knowledge, and in particular, does not measure well those relevant attributes on which the group in question happens to score well. If the item pool in question measures these attributes at all a selection routine using this group might be expected to increase the importance of these attributes in determining the total score and thereby reduce the disadvantage of the group. Therefore, the three minority groups considered here might be expected to do relatively better on the half-tests than they did on the original whole test. Their whole and half-test mean scores can be found in Tables 40 through 57 in the Appendix. Table 11 shows the frequency of such relative improvements for the four pairs of groups. As before, cases showing no differences were tabulated against the hypothesis of bias. The results parallel those for the KR 20's, with evidence of consistent improvement in the upper grades but inconsistent data from Grades 1 and 3. As was the case for proportions of unique items, the SWR (V) group does not fit the pattern.

The sets of unique items are also relatively easier for the minority groups in most cases (see Table 12); the unique-item tests are clearly biased in favor of the group used as the basis for selection and this result is true for all groups at all grades.

CONCLUSIONS

The four analyses of the data described above permit the following conclusions:

(1) Different tryout samples lead to the selection of somewhat different sets of items. Considering the restriction on range and variety of points of view represented in the item pool used, the 30% proportion of unique items, which was the average found in this study, seems large. That is, it seems likely that a majority of unique items would have been selected if the item pool had been more heterogeneous.

(2) The more economically dissimilar contrast groups are the less likely it is they will produce data leading to the selection of the same set of items.

(3) If a biased test is a test that contains a substantial proportion of items that would not have been selected had some other particular group been the tryout sample, then probably most tests are biased against most groups.

Table 11

FREQUENCIES OF CHANGES, WHOLE TEST TO HALF-TEST, IN MEAN DIFFERENCES SHOWING RELATIVELY HIGHER SCORES FOR THE COMPARISON GROUP

	Comparison Groups											
Grade	II & I		IV & III		IV & V		VI & VII		All Pairs		χ^2	p
1.6	7	1	1	8	2	7	7	2	17	18	0.1	NS
3.6	2	7	8	1	4	5	4	5	18	17	0.1	NS
5.6	7	1	8	1	1	8	8	1	24	11	4.8	.05
8.6	8	0	6	3	6	3	6	3	32	9	12.6	.001
10.6	--	--	6	3	7	2	--	--	13	5	3.6	NS
All Grades	24	9	29	16	20	25	25	11	104	60	11.8	.001
χ^2		6.8		3.8		0.6		5.4		11.8		
p		.01		.05		NS		.02		.001		

Table 12

FREQUENCY OF MEAN DIFFERENCES ON THE MINORITY GROUP UNIQUE ITEM TESTS WHICH FAVOR EACH COMPARISON GROUP WHEN CONTRASTED WITH MEAN DIFFERENCES ON THE MAJORITY GROUP UNIQUE ITEM TEST[a]

	Comparison Groups											
Grade	II & I		IV & III		IV & V		VI & VII		All Pairs		χ^2	p
1.6	5	3	6	3	8	1	8	1	27	8	10.3	.01
3.6	5	4	5	4	3	6	7	2	20	16	0.4	NS
5.6	7	1	5	4	7	2	7	2	26	9	8.3	.01
8.6	8	0	9	0	6	3	5	4	28	7	12.6	.001
10.6	--	--	6	3	5	4	--	--	11	7	0.9	NS
All Grades	25	8	31	14	29	16	27	9	112	47		
χ^2		8.8		6.4		3.8		9.0		26.6		
p		.01		.02		.05		.01		.001		

[a]Let $\bar{X}_N$ = minority mean on minority test, $\bar{X}_M$ = majority mean on minority test $\bar{Y}_N$ = minority mean on majority test, $\bar{Y}_M$ = majority mean on majority test Then $\bar{Y}_M-\bar{Y}_N-(\bar{X}_M-\bar{X}_N)>0$ favors minority, $\bar{Y}_M-\bar{Y}_N-(\bar{X}_M-\bar{X}_N)<0$ favors majority

(4) By this criterion of bias, the tests used here are more biased against minority groups than against middle-class white children. This is probably true for most published batteries of standardized tests.

(5) The proportion of unique items is a fairly good but uneven criterion of bias since in most cases these unique item tests do measure different things. What is measured depends on which group is used for selection and which group is being tested. This conclusion is not uniformly true and varies widely according to test, grade, and tryout group.

(6) The general quality of the half-tests was improved only very slightly by the item selection procedure, presumably because all the items were already a product of an item selection procedure and because the pool is rather homogeneous in style and point of view.

(7) The half-tests did improve more over the whole test for the minority groups than for the majority groups, but this improvement is small in both kinds of groups and suggests minimal bias of this sort in these tests.

(8) The use of items particularly suited to a tryout group will improve the chances of good scores among individuals from similar groups. This outcome is particularly likely in the upper grades.

(9) The amount of relative improvement in score that a minority group can expect to gain by using tryout groups like itself is not very large. This relative improvement is most unlikely to overcome any large discrepancy between typical scores in that group and those in more favored groups.

(10) It should be possible to build tests somewhat biased in favor of any group by using a fair sample of that group for item selection data.

RECOMMENDATIONS

(1) Those engaged in test construction and publishing must consider carefully the nature of their tryout groups. Probably the use of several identifiable minority groups for separate data analyses is desirable. Experience regarding the effects of variations in tryout groups is badly needed.

(2) Also needed are studies of the effects of variation in points of view among those contributing to the item pool. Would black item writers produce items better for black children?

(3) More research should be undertaken on the relative importance and value of the various possible criteria of bias, including the possibilities not considered here.

BIBLIOGRAPHY

Angelino, H., & Shedd, C. L. An initial report of a validation study of the Davis-Eells Test of General Intelligence or Problem-Solving Ability. *Journal of Psychology*, 1955, 40, 35-38.

Brown, W. M., & Russell, R. D. Limitations of admissions testing for the disadvantaged (letter). *The Personnel and Guidance Journal*, 1964, 43 (3), 301-304.

Cardall, C., & Coffman, W. E. A method for comparing the performance of different groups on the items in a test. *College Entrance Examination Board Research and Development Reports*, 1964, RB 9.

Cleary, T. A. Test bias: prediction of grades of Negro and white students in integrated colleges. *Journal of Educational Measurement*, Summer 1968, 5, 115-124.

Cleary, T. A., & Hilton, T. L. An investigation of item bias. *Educational and Psychological Measurement*, 1968, 28, 61-75.

Coleman, J. S., et al. *Equality of Educational Opportunity*. U. S. Dept. of Health, Education, & Welfare, 1966.

Eells, K., Davis, A., Havighurst, R. J., Herrick, V. E., Tyler, R. W. *Intelligence and Cultural Differences*. Chicago: University of Chicago Press, 1951.

Einhorn, H. J., & Bass, A. R. Methodological considerations relevant to discrimination in employment testing. *Psychological Bulletin*, 1971, 75 (4), 261-269.

Gilbert, H. B. On the IQ ban. *Teachers College Record*, 1966, 67, 282-285.

Green, D. R. Biased tests. 1971 (unpublished manuscript).

Green, R. L., & Farquhar, W. W. Negro academic motivation and scholastic achievement. *Journal of Educational Psychology*, 1965, 56, 241-243.

Hewer, V. H. Are tests fair to college students from homes with low socio-economic status? *Personnel and Guidance Journal*, 1965, 43, 764-769.

Hunter, L. B., & Rogers, F. A. Testing: politics and pretense. *The Urban Review*, 1967, 2 (3), 5-6, 8, 25-26.

Jensen, A. R. How much can we boost IQ and scholastic achievement? *Harvard Educational Review*, 1969, 39 (1), 1-123.

Kelley, T. L. *Interpretation of Educational Measurements*. Yonkers, N. Y.: World Book, 1927.

Kennedy, W. A., Van De Reit, V., & White, J. C. A normative sample of intelligence and achievement of Negro elementary school children in the southeastern United States. *Monographs of the Society for Research in Child Development*, 1963, 28, No. 6.

Lesser, G. S., Fifer, G., & Clark, D. H. Mental abilities of children from different social-class and cultural groups. *Monographs of the Society for Research in Child Development*, 1965, 30 (4, Whole No. 102).

Linn, R. L., & Werts, C. E. Considerations for studies of test bias. *Journal of Educational Measurement*, Spring 1971, 8 (1), 1-4.

Lord, F. M., & Novick, M. R. *Statistical Theories of Mental Test Scores*. Reading, Mass.: Addison-Wesley, 1968.

Lorge, I. Difference or bias in tests of intelligence. In A. Anastasi (Ed.) *Testing Problems in Perspective*. Washington, D.C.: American Council on Education, 1966, 456-471.

Messick, S., & Anderson, S. Educational testing, individual development, and social responsibility. *The Counseling Psychologist*, 1970, 2 (2), 80-88.

Potthoff, R. F. Statistical aspects of the problem of biases in psychological tests. *University of North Carolina Institute of Statistics Mimeo Series*, 1966, No. 479.

Ruch, G. M. *The Objective or New-type Examination*. Chicago: Scott, Foresman and Co., 1929.

Stanley, J. C., & Porter, A. C. Correlation of Scholastic Aptitude Test scores with college grades for Negroes versus whites. *Journal of Educational Measurement*, 1967, 4, 199-218.

Temp, G. Test bias: validity of the S.A.T. for blacks and whites in thirteen integrated institutions. *College Entrance Examination Board Research and Development Reports*, 1971, RB. 712.

Terman, L. M. *The Measurement of Intelligence*. Boston: Houghton Mifflin, 1916.

Terman, L. M., & Merrill, M. A. *Stanford-Binet Intelligence Scale: Manual for the Third Revision, Form L-M*. Boston: Houghton Mifflin, 1960.

Wasserman, M. Planting pansies on the roof. *The Urban Review*, 1969, 3 (3), 30-35.

Table 13

PROPORTION OF ITEMS SELECTED WHICH ARE UNIQUE
FOR ALL POSSIBLE PAIRS OF GROUPS FOR GRADE 1.6

PAIRS	TESTS								
	Vocab.	Compr.	Read. Total	Compu.	Probs.	Math. Total	Mech.	Usage	Lang. Total
I & II	41	25	40	15	42	16	42	30	--
I & III	26	17	29	10	25	18	16	30	--
I & IV	35	58	34	20	38	25	26	50	--
I & V	41	25	40	35	25	18	32	20	--
I & VI	48	50	47	25	42	30	58	20	--
I & VII	33	25	33	5	38	36	21	30	--
II & III	30	17	29	10	46	30	32	50	--
II & IV	26	58	38	30	50	25	21	50	--
II & V	17	42	26	20	29	23	21	40	--
II & VI	30	42	26	25	33	20	37	20	--
II & VII	63	42	64	15	42	36	32	50	--
III & IV	33	58	36	25	38	25	21	30	24
III & V	24	33	26	30	38	30	32	40	32
III & VI	33	58	28	25	50	36	53	40	51
III & VII	54	42	52	10	33	30	21	10	14
IV & V	35	33	34	40	42	23	21	40	27
IV & VI	33	58	40	30	38	30	53	50	51
IV & VII	43	58	41	15	54	39	21	30	22
V & VI	22	33	21	25	29	16	42	20	30
V & VII	59	42	62	30	33	43	26	40	32
VI & VII	59	42	69	25	58	41	58	40	54

Table 14

PROPORTION OF ITEMS SELECTED WHICH ARE UNIQUE
FOR ALL POSSIBLE PAIRS OF GROUPS FOR GRADE 3.6

PAIRS	TESTS								
	Vocab.	Compr.	Read. Total	Compu.	Probs.	Math. Total	Mech.	Usage	Lang. Total
I & II	30	22	28	17	35	29	48	31	41
I & III	75	22	33	14	17	19	24	15	17
I & IV	25	26	30	25	35	29	42	31	37
I & V	40	17	19	25	35	29	30	15	30
I & VI	35	26	23	17	43	27	36	38	39
I & VII	25	26	26	19	22	24	24	23	22
II & III	55	30	40	25	43	29	52	46	39
II & IV	40	22	33	17	30	27	15	15	17
II & V	25	17	26	25	30	34	33	23	30
II & VI	30	26	26	22	22	30	21	15	15
II & VII	40	26	40	25	35	27	42	38	41
III & IV	65	26	42	28	48	32	42	46	30
III & V	50	26	37	28	48	34	33	23	26
III & VI	50	30	33	22	52	36	45	46	43
III & VII	75	17	30	17	26	24	21	31	17
IV & V	35	22	28	22	35	30	30	23	28
IV & VI	40	30	28	31	22	27	21	31	24
IV & VII	30	30	37	28	35	32	36	31	37
V & VI	30	22	21	25	30	25	33	23	37
V & VII	50	17	26	25	39	34	21	23	24
VI & VII	45	35	33	25	43	32	45	46	48

Table 15

PROPORTION OF ITEMS SELECTED WHICH ARE UNIQUE
FOR ALL POSSIBLE PAIRS OF GROUPS FOR GRADE 5.6

PAIRS	TESTS								
	Vocab.	Compr.	Read. Total	Compu.	Probs.	Math. Total	Mech.	Usage	Lang. Total
I & II	50	48	46	41	50	44	45	33	--
I & III	35	33	29	21	40	30	28	29	--
I & IV	60	48	61	38	60	46	30	33	--
I & V	50	48	44	26	50	37	38	48	--
I & VI	70	57	66	44	65	54	45	38	--
I & VII	15	33	27	26	25	30	22	38	--
II & III	50	33	37	38	20	37	43	43	28
II & IV	30	24	29	18	25	26	20	33	13
II & V	25	24	24	21	25	22	15	33	16
II & VI	35	24	29	15	20	24	18	33	15
II & VII	45	43	49	44	35	41	48	48	25
III & IV	55	43	46	38	40	46	35	48	30
III & V	50	33	41	29	35	39	35	33	21
III & VI	60	33	46	44	40	48	40	33	30
III & VII	30	33	41	26	20	26	28	33	21
IV & V	35	29	37	21	20	20	25	38	16
IV & VI	35	19	27	15	15	17	22	33	13
IV & VII	55	48	61	41	55	44	38	48	25
V & VI	35	29	37	24	25	22	18	24	16
V & VII	50	38	61	29	45	33	43	38	25
VI & VII	70	52	61	41	55	46	53	33	26

Table 16

PROPORTION OF ITEMS SELECTED WHICH ARE UNIQUE
FOR ALL POSSIBLE PAIRS OF GROUPS FOR GRADE 8.6

PAIRS	TESTS								
	Vocab.	Compr.	Read. Total	Compu.	Probs.	Math. Total	Mech.	Usage	Lang. Total
I & II	40	22	26	25	36	29	42	36	--
I & III	35	22	30	25	28	20	36	[illegible]	--
I & IV	35	35	37	42	40	39	50	48	--
I & V	35	30	33	25	20	18	36	20	--
I & VI	35	17	26	33	24	22	39	20	--
I & VII	25	30	28	21	32	27	25	16	--
II & III	10	22	23	29	36	35	36	28	--
II & IV	10	30	21	33	24	33	25	36	--
II & V	15	26	21	17	28	22	28	24	--
II & VI	15	17	16	17	36	27	25	32	--
II & VII	55	35	37	25	48	35	44	40	--
III & IV	15	39	23	46	40	49	33	56	15
III & V	10	30	28	25	24	22	28	20	15
III & VI	10	26	19	33	20	29	28	24	8
III & VII	45	26	37	21	28	27	39	28	21
IV & V	15	30	21	29	36	35	42	32	15
IV & VI	20	30	23	29	32	29	33	40	20
IV & VII	50	52	51	54	56	49	56	44	28
V & VI	10	35	19	13	24	16	28	16	20
V & VII	40	35	42	29	32	29	36	24	23
VI & VII	45	39	44	29	28	29	39	16	18

Table 17

PROPORTION OF ITEMS SELECTED WHICH ARE UNIQUE
FOR ALL POSSIBLE PAIRS OF GROUPS FOR GRADE 10.6

PAIRS	TESTS								
	Vocab.	Compr.	Read. Total	Compu.	Probs.	Math. Total	Mech.	Usage	Lang. Total
III & IV	55	22	42	33	40	33	38	41	21
III & V	35	22	30	25	24	16	20	33	13
IV & V	40	22	30	33	32	24	35	30	19

Table 18

RELIABILITIES AND RELATIONSHIPS BETWEEN
ITEMS UNIQUE TO GROUPS I & II

Grade 1.6

Test	Group	No. of Unique Items	Columns* 1 r_{xx}	2 r_{yy}	3 r_{xy}	4 r'_{xy}	5 $1-(r'_{xy})^2$
Vocabulary	I	19	.82	.68	.34	.46	.79
	II		.89	.68	.37	.48	.77
Comprehension	I	3	.51	.01	.25	1.00	--
	II		.26	.24	.03	--	1.00
Reading Total	I	23	.83	.65	.42	.57	.67
	II		.89	.65	.38	.50	.75
Computation	I	3	.74	.45	.49	.85	.28
	II		.70	.65	.38	.57	.68
Concepts & Problems	I	10	.68	.56	.48	.78	.39
	II		.62	.26	.53	1.00	--
Math Total	I	7	.67	.56	.41	.67	.55
	II		.65	.48	.40	.71	.49
Mechanics	I	8	.79	.70	.46	.62	.62
	II		.85	.51	.41	.62	.61
Usage & Structure	I	3	.49	.32	.42	1.00	--
	II		.33	-.12	.15	.75	.43

*Column 1: r_{xx} is the KR 20 reliability for the unique test chosen for the group indicated.

*Column 2: r_{yy} is the KR 20 reliability for the unique test which was chosen for the comparison group.

*Column 3: r_{xy} is the correlation of scores on the two sets of unique items.

*Column 4: r'_{xy} is r_{xy} corrected for attenuation. With such low reliabilities, this correlation occasionally produces a figure greater than 1, a meaningless result. These are all recorded as 1.00.

*Column 5: $1-(r'_{xy})^2$ is an estimate of the minimum proportion of the variance which is independent.

Table 19

RELIABILITIES AND RELATIONSHIPS BETWEEN
ITEMS UNIQUE TO GROUPS I & II

Grade 3.6

Test	Group	No. of Unique Items	Columns* 1 r_{xx}	2 r_{yy}	3 r_{xy}	4 r'_{xy}	5 $1-(r'_{xy})^2$
Vocabulary	I	6	.82	.82	.40	.49	.76
	II		.80	.59	.46	.67	.55
Comprehension	I	5	.73	.62	.69	1.00	--
	II		.67	.54	.55	.92	.15
Reading Total	I	12	.86	.74	.79	.99	.02
	II		.83	.67	.64	.86	.26
Computation	I	6	.66	.45	.45	.83	.31
	II		.83	.66	.50	.68	.54
Concepts & Problems	I	8	.78	.39	.64	1.00	--
	II		.70	.50	.49	.83	.31
Math Total	I	17	.87	.63	.72	.97	.06
	II		.84	.74	.63	.80	.36
Mechanics	I	16	.90	.82	.62	.72	.48
	II		.84	.66	.58	.78	.39
Usage & Structure	I	4	.63	.48	.41	.75	.44
	II		.62	.38	.33	.67	.55
Language Total	I	19	.90	.82	.65	.76	.42
	II		.88	.60	.53	.74	.45

*See footnotes for Table 18.

Table 20

RELIABILITIES AND RELATIONSHIPS BETWEEN
ITEMS UNIQUE TO GROUPS I & II

Grade 5.6

Test	Group	No. of Unique Items	Columns* 1 r_{xx}	2 r_{yy}	3 r_{xy}	4 r'_{xy}	5 $1-(r'_{xy})^2$
Vocabulary	I	10	.83	.54	.49	.73	.46
	II		.78	.51	.25	.40	.84
Comprehension	I	10	.71	.58	.55	.85	.27
	II		.66	.35	.36	.75	.44
Reading Total	I	19	.86	.68	.60	.79	.38
	II		.81	.54	.34	.51	.74
Computation	I	14	.79	.55	.35	.53	.72
	II		.83	.61	.39	.55	.70
Concepts & Problems	I	10	.69	.58	.35	.56	.69
	II		.69	.42	.39	.71	.49
Math Total	I	24	.84	.66	.47	.63	.60
	II		.86	.64	.51	.69	.53
Mechanics	I	18	.85	.77	.50	.61	.62
	II		.84	.61	.48	.67	.55
Usage & Structure	I	7	.36	.03	.12	1.00	--
	II		.36	.12	.20	.96	.07

*See footnotes for Table 18.

Table 21

RELIABILITIES AND RELATIONSHIPS BETWEEN
ITEMS UNIQUE TO GROUPS I & II

Grade 8.6

Test	Group	No. of Unique Items	Columns* 1 r_{xx}	2 r_{yy}	3 r_{xy}	4 r'_{xy}	5 $1-(r'_{xy})^2$
Vocabulary	I	8	.65	.54	.57	.96	.07
	II		.73	.50	.54	.89	.20
Comprehension	I	5	.50	.34	.43	1.00	--
	II		.46	.36	.27	.66	.56
Reading Total	I	11	.72	.56	.51	.81	.35
	II		.74	.54	.43	.68	.54
Computation	I	6	.67	.50	.51	.88	.22
	II		.69	.51	.44	.74	.45
Concepts & Problems	I	9	.70	.52	.51	.84	.29
	II		.62	.38	.44	.91	.18
Math Total	I	14	.76	.65	.66	.94	.12
	II		.74	.56	.54	.84	.30
Mechanics	I	15	.80	.65	.47	.65	.58
	II		.85	.74	.58	.73	.47
Usage & Structure	I	9	.49	.11	.19	.82	.33
	II		.51	.33	.16	.39	.85

*See footnotes for Table 18.

Table 22

RELIABILITIES AND RELATIONSHIPS BETWEEN
ITEMS UNIQUE TO GROUPS III & IV

Grade 1.6

Test	Group	No. of Unique Items	Columns* 1 r_{xx}	2 r_{yy}	3 r_{xy}	4 r'_{xy}	5 $1-(r'_{xy})^2$
Vocabulary	III	15	.76	.61	.54	.79	.37
	IV		.88	.78	.77	.93	.14
Comprehension	III	7	.58	.21	.09	.26	.93
	IV		.81	.74	.72	.93	.13
Reading Total	III	21	.75	.58	.50	.76	.42
	IV		.91	.83	.78	.90	.19
Computation	III	5	.73	.56	.57	.89	.20
	IV		.79	.74	.77	1.00	--
Concepts & Problems	III	9	.64	.49	.54	.96	.07
	IV		.82	.70	.76	1.00	--
Math Total	III	11	.72	.50	.62	1.00	--
	IV		.86	.77	.79	.97	.06
Mechanics	III	4	.58	.57	.33	.57	.67
	IV		.81	.49	.56	.89	.21
Usage & Structure	III	3	.48	.15	.27	1.00	--
	IV		.70	.46	.60	1.00	--
Language Total	III	9	.67	.43	.56	1.00	--
	IV		.85	.73	.79	1.00	--

*See footnotes for Table 18.

Table 23

RELIABILITIES AND RELATIONSHIPS BETWEEN
ITEMS UNIQUE TO GROUPS III & IV

Grade 3.6

Test	Group	No. of Unique Items	Columns* 1 r_{xx}	2 r_{yy}	3 r_{xy}	4 r'_{xy}	5 $1-(r'_{xy})^2$
Vocabulary	III	13	.77	.73	.35	.47	.78
	IV		.86	.68	.57	.75	.44
Comprehension	III	6	.75	.40	.42	.77	.41
	IV		.72	.43	.61	1.00	--
Reading Total	III	18	.83	.57	.57	.83	.21
	IV		.87	.71	.76	.97	.06
Computation	III	10	.87	.60	.39	.54	.71
	IV		.88	.78	.43	.52	.73
Concepts & Problems	III	11	.68	.23	.42	1.00	--
	IV		.85	.70	.76	.98	.03
Math Total	III	19	.85	.49	.60	.93	.14
	IV		.89	.79	.68	.81	.34
Mechanics	III	14	.85	.71	.67	.86	.26
	IV		.87	.77	.59	.72	.48
Usage & Structure	III	6	.60	.43	.34	.67	.55
	IV		.68	.34	.33	.68	.53
Language Total	III	14	.83	.66	.54	.73	.47
	IV		.85	.68	.48	.63	.60

*See footnotes for Table 18.

Table 24

RELIABILITIES AND RELATIONSHIPS BETWEEN
ITEMS UNIQUE TO GROUPS III & IV

Grade 5.6

Test	Group	No. of Unique Items	Columns* 1 r_{xx}	2 r_{yy}	3 r_{xy}	4 r'_{xy}	5 $1-(r'_{xy})^2$
Vocabulary	III	11	.76	.49	.65	1.00	--
	IV		.81	.52	.51	.79	.38
Comprehension	III	9	.70	.51	.60	1.00	--
	IV		.70	.50	.40	.68	.54
Reading Total	III	19	.83	.66	.72	.97	.05
	IV		.83	.63	.57	.79	.38
Computation	III	13	.78	.66	.43	.60	.64
	IV		.84	.49	.42	.66	.57
Concepts & Problems	III	8	.66	.48	.39	.61	.63
	IV		.77	.53	.48	.75	.44
Math Total	III	25	.85	.71	.60	.77	.40
	IV		.90	.70	.59	.74	.45
Mechanics	III	14	.80	.69	.53	.71	.49
	IV		.82	.49	.62	.98	.04
Usage & Structure	III	10	.53	.16	.32	1.00	--
	IV		.55	.02	.17	1.00	--
Language Total	III	18	.80	.56	.58	1.00	--
	IV		.78	.37	.48	.89	.20

*See footnotes for Table 18.

Table 25

RELIABILITIES AND RELATIONSHIPS BETWEEN
ITEMS UNIQUE TO GROUPS III & IV

Grade 8.6

Test	Group	No. of Unique Items	Columns* 1 r_{xx}	2 r_{yy}	3 r_{xy}	4 r'_{xy}	5 $1-(r'_{xy})^2$
Vocabulary	III	3	.50	.58	.49	.91	.17
	IV		.46	.22	.33	1.00	--
Comprehension	III	9	.70	.60	.59	.91	.17
	IV		.61	.20	.29	.83	.31
Reading Total	III	10	.74	.61	.65	.97	.06
	IV		.62	.36	.46	.98	.05
Computation	III	11	.82	.70	.51	.72	.48
	IV		.72	.45	.20	.35	.88
Concepts & Problems	III	10	.77	.60	.62	.91	.17
	IV		.64	.40	.49	.97	.06
Math Total	III	24	.89	.80	.68	.81	.45
	IV		.83	.56	.50	.73	.46
Mechanics	III	12	.85	.69	.64	.84	.30
	IV		.80	.72	.66	.87	.24
Usage & Structure	III	8	.44	.03	.18	1.00	--
	IV		.30	.14	.08	.39	.85
Language Total	III	9	.75	.42	.46	.82	.33
	IV		.67	.41	.43	.82	.33

*See footnotes for Table 18.

Table 26

RELIABILITIES AND RELATIONSHIPS BETWEEN
ITEMS UNIQUE TO GROUPS III & V

Grade 10.6

Test	Group	No. of Unique Items	Columns* 1 r_{xx}	2 r_{yy}	3 r_{xy}	4 r'_{xy}	5 $1-(r'_{xy})^2$
Vocabulary	III	11	.80	.57	.63	.93	.13
	IV		.64	.36	.35	.76	.47
Comprehension	III	5	.55	.41	.40	.84	.29
	IV		.46	.02	.19	1.00	--
Reading Total	III	18	.85	.64	.70	.95	.10
	IV		.74	.32	.34	.70	.51
Computation	III	8	.69	.57	.58	.93	.14
	IV		.80	.68	.66	.89	.20
Concepts & Problems	III	10	.74	.48	.58	.97	.05
	IV		.75	.48	.36	.60	.64
Math Total	III	16	.81	.71	.64	.84	.29
	IV		.87	.70	.59	.75	.43
Mechanics	III	15	.83	.68	.62	.84	.30
	IV		.90	.78	.68	.81	.34
Usage & Structure	III	11	.51	.18	.35	1.00	--
	IV		.35	.24	.07	.25	.94
Language Total	III	14	.81	.64	.39	.54	.71
	IV		.86	.49	.40	.88	.62

*See footnotes for Table 18.

Table 27

RELIABILITIES AND RELATIONSHIPS BETWEEN
ITEMS UNIQUE TO GROUPS IV & V

Grade 1.6

Test	Group	No. of Unique Items	Columns* 1 r_{xx}	2 r_{yy}	3 r_{xy}	4 r'_{xy}	5 $1-(r'_{xy})^2$
Vocabulary	IV	16	.90	.83	.67	.77	.40
	V		.87	.71	.63	.80	.36
Comprehension	IV	4	.73	.62	.50	.74	.45
	V		.43	.38	.30	.74	.45
Reading Total	IV	20	.91	.86	.67	.75	.43
	V		.89	.70	.65	.82	.32
Computation	IV	8	.88	.83	.73	.85	.27
	V		.88	.81	.67	.79	.37
Concepts & Problems	IV	10	.81	.68	.77	1.00	--
	V		.80	.61	.68	.97	.05
Math Total	IV	10	.85	.74	.76	.96	.08
	V		.81	.68	.69	.93	.14
Mechanics	IV	4	.83	.68	.69	.92	.16
	V		.43	.51	.24	.51	.74
Usage & Structure	IV	4	.70	.65	.57	.84	.29
	V		.66	.46	.39	.71	.50
Language Total	IV	10	.89	.73	.69	.85	.27
	V		.73	.43	.42	.75	.44

*See footnotes for Table 18.

Table 28

RELIABILITIES AND RELATIONSHIPS BETWEEN
ITEMS UNIQUE TO GROUPS IV & V

Grade 3.6

Test	Group	No. of Unique Items	Columns* 1 r_{xx}	2 r_{yy}	3 r_{xy}	4 r'_{xy}	5 $1-(r'_{xy})^2$
Vocabulary	IV	7	.71	.60	.54	.83	.31
	V		.79	.68	.62	.85	.28
Comprehension	IV	5	.68	.49	.61	1.00	--
	V		.70	.54	.55	.89	.20
Reading Total	IV	12	.81	.69	.71	.95	.10
	V		.83	.68	.74	.98	.03
Computation	IV	8	.76	.67	.63	.88	.22
	V		.76	.65	.68	.97	.06
Concepts & Problems	IV	8	.76	.61	.71	.97	.06
	V		.77	.61	.67	.98	.04
Math Total	IV	18	.87	.77	.82	1.00	--
	V		.86	.75	.81	1.00	--
Mechanics	IV	10	.81	.68	.53	.71	.49
	V		.79	.76	.59	.76	.42
Usage & Structure	IV	3	.46	.12	.25	1.00	--
	V		.50	.22	.49	1.00	--
Language Total	IV	13	.82	.65	.57	.78	.39
	V		.80	.72	.70	.92	.15

*See footnotes for Table 18.

Table 29

RELIABILITIES AND RELATIONSHIPS BETWEEN
ITEMS UNIQUE TO GROUPS IV & V

Grade 5.6

Test	Group	No. of Unique Items	Columns* 1 r_{xx}	2 r_{yy}	3 r_{xy}	4 r'_{xy}	5 $1-(r'_{xy})^2$
Vocabulary	IV	7	.63	.42	.48	.93	.13
	V		.72	.55	.61	.97	.06
Comprehension	IV	6	.64	.34	.41	.88	.23
	V		.60	.42	.52	1.00	--
Reading Total	IV	15	.77	.60	.65	.95	.09
	V		.81	.70	.73	.98	.04
Computation	IV	7	.74	.47	.41	.69	.52
	V		.73	.59	.42	.64	.59
Concepts & Problems	IV	4	.62	.36	.37	.78	.39
	V		.53	.48	.46	.91	.17
Math Total	IV	11	.76	.65	.58	.82	.32
	V		.82	.66	.58	.79	.38
Mechanics	IV	10	.76	.59	.57	.85	.28
	V		.79	.69	.62	.84	.29
Usage & Structure	IV	8	.51	.11	.08	.33	.89
	V		.44	.06	-.02	--	.98
Language Total	IV	10	.57	.40	.42	.88	.23
	V		.70	.48	.61	1.00	--

*See footnotes for Table 18.

Table 30

RELIABILITIES AND RELATIONSHIPS BETWEEN
ITEMS UNIQUE TO GROUPS IV & V

Grade 8.6

Test	Group	No. of Unique Items	Columns* 1 r_{xx}	2 r_{yy}	3 r_{xy}	4 r'_{xy}	5 $1-(r'_{xy})^2$
Vocabulary	IV	3	.46	.28	.33	.92	.15
	V		.54	.43	.44	.91	.17
Comprehension	IV	7	.52	.22	.30	.89	.21
	V		.61	.50	.35	.63	.60
Reading Total	IV	9	.59	.32	.47	1.00	--
	V		.67	.54	.55	.92	.16
Computation	IV	7	.60	.25	.23	.59	.65
	V		.74	.61	.54	.81	.35
Concepts & Problems	IV	9	.66	.34	.44	.93	.14
	V		.70	.45	.57	1.00	--
Math Total	IV	17	.79	.44	.42	.71	.49
	V		.82	.71	.65	.85	.27
Mechanics	IV	15	.85	.81	.58	.70	.51
	V		.86	.72	.60	.76	.42
Usage & Structure	IV	8	.34	.18	-.04	.17	.97
	V		.50	.10	.14	.62	.61
Language Total	IV	9	.68	.41	.37	.70	.51
	V		.64	.45	.40	.75	.44

*See footnotes for Table 18.

Table 31

RELIABILITIES AND RELATIONSHIPS BETWEEN
ITEMS UNIQUE TO GROUPS IV & V

Grade 10.6

Test	Group	No. of Unique Items	Columns* 1 r_{xx}	2 r_{yy}	3 r_{xy}	4 r'_{xy}	5 $1-(r'_{xy})^2$
Vocabulary	IV	8	.63	.15	.37	1.00	--
	V		.63	.49	.47	.85	.28
Comprehension	IV	5	.29	.05	.02	.17	.92
	V		.47	.25	.28	.79	.38
Reading Total	IV	13	.62	.25	.44	1.00	--
	V		.72	.51	.56	.92	.15
Computation	IV	8	.80	.69	.69	.93	.14
	V		.80	.61	.67	.96	.08
Concepts & Problems	IV	8	.66	.41	.27	.52	.73
	V		.69	.39	.44	.85	.28
Math Total	IV	12	.82	.63	.56	.78	.39
	V		.78	.62	.58	.84	.30
Mechanics	IV	14	.90	.78	.58	.69	.52
	V		.86	.75	.54	.67	.55
Usage & Structure	IV	8	.25	-.13	.09	.50	.75
	V		.39	-.003	.29	--	--
Language Total	IV	13	.86	.47	.43	.68	.54
	V		.79	.68	.43	.58	.66

*See footnotes for Table 18.

Table 32

RELIABILITIES AND RELATIONSHIPS BETWEEN
ITEMS UNIQUE TO GROUPS VI & VII

Grade 1.6

Test	Group	No. of Unique Items	Columns* 1 r_{xx}	2 r_{yy}	3 r_{xy}	4 r'_{xy}	5 $1-(r'_{xy})^2$
Vocabulary	VI	27	.88	.37	.34	.60	.64
	VII		:90	.67	.46	.59	.65
Comprehension	VI	5	.18	-.01	.10	1.00	--
	VII		.76	.52	.61	.97	.06
Reading Total	VI	40	.89	.44	.20	.32	.90
	VII		.94	.81	.54	.62	.62
Computation	VI	5	.70	.73	.32	.45	.80
	VII		.90	.56	.41	.57	.67
Concepts & Problems	VI	14	.77	.35	.47	.91	.18
	VII		.81	.63	.68	.95	.09
Math Total	VI	18	.86	.54	.58	.85	.28
	VII		.85	.70	.67	.86	.25
Mechanics	VI	11	.74	.45	-.17	.30	.91
	VII		.92	.81	.64	.74	.45
Usage & Strucutre	VI	4	.56	.45	.19	.37	.86
	VII		.56	.42	.50	1.00	--
Language Total	VI	20	.73	.40	.15	.28	.92
	VII		.92	.67	.60	.76	.42

*See footnotes for Table 18.

Table 33

RELIABILITIES AND RELATIONSHIPS BETWEEN
ITEMS UNIQUE TO GROUPS VI & VII

Grade 3.6

Test	Group	No. of Unique Items	Columns* 1 r_{xx}	2 r_{yy}	3 r_{xy}	4 r'_{xy}	5 $1-(r'_{xy})^2$
Vocabulary	VI	9	.81	.59	.53	.77	.41
	VII		.85	.56	.47	.68	.54
Comprehension	VI	8	.76	.54	.61	.95	.09
	VII		.75	.59	.64	.96	.07
Reading Total	VI	14	.85	.65	.53	.71	.49
	VII		.83	.61	.65	.91	.17
Computation	VI	9	.70	.66	.50	.73	.46
	VII		.81	.48	.52	.84	.30
Concepts & Problems	VI	10	.80	.53	.64	.98	.03
	VII		.65	.38	.44	.88	.22
Math Total	VI	19	.86	.74	.52	.65	.58
	VII		.85	.63	.61	.84	.30
Mechanics	VI	15	.87	.50	.54	.82	.33
	VII		.85	.72	.63	.81	.35
Usage & Strucutre	VI	6	.72	.47	.22	.37	.86
	VII		.61	.37	.37	.78	.39
Language Total	VI	22	.90	.61	.43	.58	.66
	VII		.88	.73	.60	.74	.45

*See footnotes for Table 18.

Table 34

RELIABILITIES AND RELATIONSHIPS BETWEEN
ITEMS UNIQUE TO GROUPS VI & VII

Grade 5.6

Test	Group	No. of Unique Items	Columns* 1 r_{xx}	2 r_{yy}	3 r_{xy}	4 r'_{xy}	5 $1-(r'_{xy})^2$
Vocabulary	VI	14	.83	.61	.44	.62	.62
	VII		.83	.62	.66	.92	.15
Comprehension	VI	11	.75	.40	.32	.58	.66
	VII		.73	.59	.53	.81	.35
Reading Total	VI	25	.87	.62	.51	.69	.52
	VII		.86	.70	.73	.94	.11
Computation	VI	14	.83	.53	.37	.56	.69
	VII		.75	.63	.45	.66	.57
Concepts & Problems	VI	11	.81	.45	.33	.55	.70
	VII		.76	.38	.34	.63	.60
Math Total	VI	25	.89	.70	.35	.45	.80
	VII		.84	.63	.45	.62	.62
Mechanics	VI	21	.85	.55	.46	.67	.55
	VII		.87	.75	.59	.73	.47
Usage & Structure	VI	7	.25	.37	-.08	.26	.93
	VII		.42	.21	.17	.57	.67
Language Total	VI	17	.74	.47	.21	.36	.87
	VII		.83	.53	.48	.72	.48

*See footnotes for Table 18.

Table 35

RELIABILITIES AND RELATIONSHIPS BETWEEN
ITEMS UNIQUE TO GROUPS VI & VII

Grade 8.6

Test	Group	No. of Unique Items	Columns* 1 r_{xx}	2 r_{yy}	3 r_{xy}	4 r'_{xy}	5 $1-(r'_{xy})^2$
Vocabulary	VI	9	.76	.57	.42	.64	.59
	VII		.75	.65	.44	.63	.60
Comprehension	VI	9	.65	.45	.49	.91	.18
	VII		.71	.56	.61	.97	.06
Reading Total	VI	19	.84	.67	.58	.77	.40
	VII		.84	.73	.62	.79	.37
Computation	VI	7	.77	.66	.39	.55	.70
	VII		.77	.61	.45	.66	.57
Concepts & Problems	VI	7	.70	.40	.39	.73	.46
	VII		.63	.35	.40	.85	.27
Math Total	VI	14	.82	.62	.50	.70	.51
	VII		.73	.59	.62	.94	.11
Mechanics	VI	14	.81	.72	.63	.82	.32
	VII		.84	.70	.61	.79	.37
Usage & Structure	VI	4	.94	.09	.06	.20	.96
	VII		.31	.10	.20	1.00	--
Language Total	VI	11	.74	.52	.33	.53	.72
	VII		.73	.61	.39	.58	.66

*See footnotes for Table 18.

Table 36

MEANS AND CROSSOVER MEANS ON UNIQUE ITEM TESTS
FOR GROUPS I & II

		GRADE 1.6			GRADE 3.6		
TEST	GROUP	No. of Unique Items	Mean	Cross-over Mean[a]	No. of Unique Items	Mean	Cross-over Mean[a]
Vocabulary	I	19	10.5	17.1	6	5.3	5.5
	II		13.7	6.6		3.5	4.0
Comprehension	I	3	0.7	0.9	5	4.0	3.8
	II		0.7	0.4		2.4	2.5
Reading Total	I	23	9.5	19.5	12	9.4	10.3
	II		15.4	6.5		7.0	5.1
Computation	I	3	1.4	2.3	6	3.3	5.4
	II		1.9	0.9		4.5	2.5
Concepts & Problems	I	10	5.8	7.9	8	6.0	7.0
	II		5.4	3.3		5.2	3.2
Math Total	I	7	3.7	6.0	17	13.2	13.2
	II		3.8	2.4		10.0	8.4
Mechanics	I	8	3.2	5.3	16	6.2	11.6
	II		3.5	1.5		6.1	2.0
Usage & Structure	I	3	2.0	1.8	4	2.7	3.6
	II		0.9	0.7		2.6	1.4
Language Total	I	--	- -	- -	19	8.4	15.2
	II		- -	- -		9.4	3.2

[a]A crossover mean is the mean of the group on the unique item test selected for the group with which it is being compared.

Table 36 (Continued)

MEANS AND CROSSOVER MEANS ON UNIQUE ITEM TESTS
FOR GROUPS I & II

		GRADE 5.6			GRADE 8.6		
TEST	GROUP	No. of Unique Items	Mean	Cross-over Mean[a]	No. of Unique Items	Mean	Cross-over Mean[a]
Vocabulary	I	10	6.1	9.0	8	5.2	6.9
	II		6.4	2.7		5.5	3.8
Comprehension	I	10	6.2	8.7	5	2.5	3.7
	II		6.1	3.5		2.7	1.6
Reading Total	I	19	11.3	16.9	11	6.6	9.3
	II		12.1	5.7		7.5	4.7
Computation	I	14	7.5	12.4	6	2.7	3.4
	II		10.8	4.2		2.6	1.9
Concepts & Problems	I	10	4.5	8.9	9	4.1	6.5
	II		6.6	1.7		5.3	2.5
Math Total	I	24	13.2	21.0	14	6.6	10.1
	II		17.6	7.0		8.1	4.3
Mechanics	I	18	8.5	14.6	15	5.9	11.7
	II		9.3	3.6		9.1	3.1
Usage & Structure	I	7	3.9	4.3	9	3.9	6.1
	II		3.7	2.8		5.7	2.5
Language Total	I	--	--	--	--	--	--
	II		--	--		--	--

[a]A crossover mean is the mean of the group on the unique item test selected for the group with which it is being compared.

Table 37

MEANS AND CROSSOVER MEANS ON UNIQUE ITEM TESTS
FOR GROUPS III & IV

		GRADE 1.6			GRADE 3.6		
TEST	GROUP	No. of Unique Items	Mean	Cross-over Mean[a]	No. of Unique Items	Mean	Cross-over Mean[a]
Vocabulary	III	15	13.3	8.8	13	11.0	12.7
	IV		7.7	10.5		9.0	6.5
Comprehension	III	7	3.4	1.6	6	4.9	4.9
	IV		2.3	3.3		2.7	2.8
Reading Total	III	21	16.9	8.9	18	14.9	16.5
	IV		8.7	13.8		10.6	9.1
Computation	III	5	4.0	3.7	10	7.0	9.5
	IV		2.1	2.1		7.3	4.5
Concepts & Problems	III	9	5.3	6.6	11	9.0	10.5
	IV		4.6	4.3		7.8	6.2
Math Total	III	11	6.3	7.7	19	14.4	17.9
	IV		5.1	4.7		13.4	9.7
Mechanics	III	4	1.7	2.4	14	5.2	10.9
	IV		1.4	1.6		5.1	1.4
Usage & Structure	III	3	2.3	2.0	6	4.2	5.4
	IV		1.4	1.2		3.4	1.7
Language Total	III	9	4.5	3.3	14	5.9	12.6
	IV		2.6	3.3		7.5	1.8

[a]A crossover mean is the mean of the group on the unique item test selected for the group with which it is being compared.

Table 37 (Continued)

MEANS AND CROSSOVER MEANS ON UNIQUE ITEM TESTS
FOR GROUPS III & IV

		GRADE 5.6			GRADE 8.6		
TEST	GROUP	No. of Unique Items	Mean	Cross-over Mean[a]	No. of Unique Items	Mean	Cross-over Mean[a]
Vocabulary	III	11	7.8	9.6	3	2.3	2.6
	IV		5.5	3.4		1.7	1.3
Comprehension	III	9	6.5	7.7	9	5.0	7.0
	IV		4.5	2.8		5.5	3.0
Reading Total	III	19	13.6	16.5	10	6.3	7.9
	IV		9.2	5.9		5.1	3.5
Computation	III	13	5.9	11.6	11	5.7	8.3
	IV		8.4	3.0		4.5	1.4
Concepts & Problems	III	8	5.3	7.2	10	5.9	7.5
	IV		4.8	2.4		5.2	3.2
Math Total	III	25	13.0	22.0	24	13.0	18.6
	IV		14.8	6.2		12.3	5.0
Mechanics	III	14	6.8	10.8	12	5.4	8.3
	IV		6.2	2.9		5.6	2.4
Usage & Structure	III	10	7.0	4.8	8	4.1	4.2
	IV		4.0	5.1		4.2	2.6
Language Total	III	18	7.5	13.5	9	4.2	5.6
	IV		9.2	3.5		4.7	1.8

[a] A crossover mean is the mean of the group on the unique item test selected for the group with which it is being compared.

Table 37 (Continued)

MEANS AND CROSSOVER MEANS ON UNIQUE ITEM TESTS
FOR GROUPS III & IV

GRADE 10.6				
TEST	GROUP	No. of Unique Items	Mean	Cross-over Mean[a]
Vocabulary	III	11	5.9	9.2
	IV		5.8	2.9
Comprehension	III	5	3.0	3.6
	IV		2.4	1.6
Reading Total	III	18	9.2	14.1
	IV		9.9	4.4
Computation	III	8	5.4	5.9
	IV		3.6	2.8
Concepts & Problems	III	10	4.7	7.7
	IV		4.8	1.9
Math Total	III	16	8.1	12.2
	IV		7.1	3.6
Mechanics	III	15	9.2	12.9
	IV		9.3	4.3
Usage & Structure	III	11	5.6	6.6
	IV		5.2	3.7
Language Total	III	14	7.0	12.4
	IV		9.8	3.4

[a]A crossover mean is the mean of the group on the unique item test selected for the group with which it is being compared.

Table 38

MEANS AND CROSSOVER MEANS ON UNIQUE ITEM TESTS
FOR GROUPS IV & V

		GRADE 1.6			GRADE 3.6		
TEST	GROUP	No. of Unique Items	Mean	Cross-over Mean[a]	No. of Unique Items	Mean	Cross-over Mean[a]
Vocabulary	IV	16	7.1	12.2	7	4.5	3.9
	V		12.0	6.7		5.2	5.8
Comprehension	IV	4	1.3	1.7	5	2.3	2.4
	V		1.3	0.9		3.6	3.6
Reading Total	IV	20	8.5	14.8	12	6.6	6.2
	V		14.1	7.5		8.6	9.1
Computation	IV	8	3.1	4.1	8	5.1	4.1
	V		5.0	3.5		5.0	6.0
Concepts & Problems	IV	10	4.4	5.4	8	4.8	4.8
	V		5.8	4.9		5.7	5.9
Math Total	IV	10	4.5	5.5	18	11.3	10.5
	V		6.3	4.5		12.6	13.5
Mechanics	IV	4	1.3	1.5	10	4.0	1.2
	V		1.7	0.9		2.8	6.1
Usage & Structure	IV	4	1.4	2.3	3	1.6	0.8
	V		2.4	1.4		1.6	2.2
Language Total	IV	10	2.8	5.7	13	5.7	2.5
	V		5.7	2.0		4.6	8.5

[a]A crossover mean is the mean of the group on the unique item test selected for the group with which it is being compared.

Table 38 (Continued)

MEANS AND CROSSOVER MEANS ON UNIQUE ITEM TESTS
FOR GROUPS IV & V

		GRADE 5.6			GRADE 8.6		
TEST	GROUP	No. of Unique Items	Mean	Cross-over Mean[a]	No. of Unique Items	Mean	Cross-over Mean[a]
Vocabulary	IV	7	2.6	2.4	3	1.7	1.0
	V		3.9	3.7		1.6	2.4
Comprehension	IV	6	2.9	1.7	7	4.3	2.6
	V		2.7	4.1		3.3	5.0
Reading Total	IV	15	6.7	5.2	9	5.6	3.7
	V		7.7	9.4		4.9	7.2
Computation	IV	7	4.7	2.6	7	3.8	1.1
	V		4.1	5.7		3.5	5.1
Concepts & Problems	IV	4	1.9	1.3	9	4.9	2.7
	V		2.4	2.7		4.5	6.4
Math Total	IV	11	6.2	3.9	17	9.6	3.8
	V		6.5	8.4		8.3	12.6
Mechanics	IV	10	3.7	4.3	15	7.8	2.8
	V		6.8	6.0		6.1	10.2
Usage & Structure	IV	8	3.2	4.2	8	4.7	2.3
	V		5.1	3.3		3.2	5.0
Language Total	IV	10	4.6	2.5	9	5.1	1.5
	V		4.5	5.5		3.0	6.2

[a]A crossover mean is the mean of the group on the unique item test selected for the group with which it is being compared.

Table 38 (Continued)

MEANS AND CROSSOVER MEANS ON UNIQUE ITEM TESTS
FOR GROUPS IV & V

		GRADE 10.6		
TEST	GROUP	No. of Unique Items	Mean	Cross-over Mean[a]
Vocabulary	IV	8	4.5	2.5
	V		3.5	6.0
Comprehension	IV	5	2.3	1.8
	V		2.5	2.8
Reading Total	IV	13	6.8	4.2
	V		6.0	8.5
Computation	IV	8	3.5	2.3
	V		4.7	5.9
Concepts & Problems	IV	8	4.0	1.7
	V		2.7	5.9
Math Total	IV	12	5.5	2.8
	V		5.0	8.9
Mechanics	IV	14	9.1	4.3
	V		7.5	11.8
Usage & Structure	IV	8	3.8	3.5
	V		4.6	4.1
Language Total	IV	13	8.9	2.5
	V		4.6	11.0

[a]A crossover mean is the mean of the group on the unique item test selected for the group with which it is being compared.

Table 39

MEANS AND CROSSOVER MEANS ON UNIQUE ITEM TESTS
FOR GROUPS VI & VII

		GRADE 1.6			GRADE 3.6		
TEST	GROUP	No. of Unique Items	Mean	Cross-over Mean[a]	No. of Unique Items	Mean	Cross-over Mean[a]
Vocabulary	VI	27	16.5	7.1	9	3.9	6.3
	VII		16.6	25.7		8.4	8.3
Comprehension	VI	5	1.2	0.9	8	3.4	3.2
	VII		2.4	2.1		6.7	6.8
Reading Total	VI	40	23.5	9.6	14	6.9	5.3
	VII		21.4	37.5		11.6	12.6
Computation	VI	5	2.6	1.1	9	5.9	3.6
	VII		3.4	4.3		7.4	8.3
Concepts & Problems	VI	14	6.0	4.1	10	5.8	4.2
	VII		8.8	10.4		8.2	9.6
Math Total	VI	18	9.5	4.8	19	11.1	9.0
	VII		11.2	15.3		16.0	17.8
Mechanics	VI	11	3.0	1.7	15	6.9	3.1
	VII		5.8	6.6		6.7	10.3
Usage & Structure	VI	4	1.7	1.0	6	3.3	1.5
	VII		2.6	3.3		3.8	5.3
Language Total	VI	20	6.4	3.2	22	11.5	4.1
	VII		10.2	13.0		9.7	17.7

[a]A crossover mean is the mean of the group on the unique item test selected for the group with which it is being compared.

Table 39 (Continued)
MEANS AND CROSSOVER MEANS ON UNIQUE ITEM TESTS
FOR GROUPS VI & VII

		GRADE 5.6			GRADE 8.6		
TEST	GROUP	No. of Unique Items	Mean	Cross-over Mean[a]	No. of Unique Items	Mean	Cross-over Mean[a]
Vocabulary	VI	14	7.7	4.1	9	5.5	2.9
	VII		9.5	12.6		5.8	8.5
Comprehension	VI	11	6.8	3.2	9	5.3	3.3
	VII		6.9	9.8		5.3	7.7
Reading Total	VI	25	15.0	7.5	19	11.4	6.3
	VII		16.5	22.3		11.3	17.5
Computation	VI	14	10.3	4.0	7	3.9	2.6
	VII		7.3	12.6		4.8	5.7
Concepts & Problems	VI	11	7.7	2.3	7	4.2	2.1
	VII		6.2	10.4		4.1	5.9
Math Total	VI	25	17.4	5.9	14	8.4	5.0
	VII		13.6	22.6		9.0	11.6
Mechanics	VI	21	10.8	3.9	14	10.1	5.2
	VII		9.6	17.1		9.0	11.6
Usage & Structure	VI	7	3.2	2.9	4	1.8	2.3
	VII		3.9	3.8		2.3	2.0
Language Total	VI	17	9.1	3.2	11	7.8	4.4
	VII		6.8	12.9		6.5	9.0

[a]A crossover mean is the mean of the group on the unique item test selected for the group with which it is being compared.

Table 40

MEANS AND STANDARD DEVIATIONS ON THE WHOLE AND HALF-TESTS FOR NORTHERN WHITE SUBURBAN VERSUS NORTHERN BLACK URBAN

Grade 1.6

Test	Group	Whole Test		Half-Test	
		Mean*	SD	Mean*	SD
Vocabulary	I	67.7	10.6	32.6	8.2
	II	48.7	15.0	26.3	10.0
Comprehension	I	7.8	4.2	4.4	2.9
	II	5.9	3.4	3.6	2.4
Reading Total	I	75.6	13.5	35.8	10.1
	II	56.5	16.8	31.8	12.2
Computation	I	22.9	10.0	10.7	6.7
	II	16.3	10.1	8.3	6.2
Concepts & Problems	I	31.5	7.5	15.7	5.1
	II	21.1	7.4	12.2	5.1
Math Total	I	54.4	15.4	26.6	10.8
	II	37.6	15.5	20.2	11.0
Mechanics	I	16.0	7.9	8.4	5.6
	II	9.6	7.5	6.7	5.4
Usage & Structure	I	12.7	3.8	7.7	2.3
	II	6.4	3.6	3.9	2.6
Language Total	I	38.8	14.7	--	--
	II	23.9	11.1	--	--

*The mean on the whole test for Reading, Mathematics, and Language Totals are equal to the sum of the means of the subtests. This is not true of the half-test means, since the items for each test were selected separately.

Table 41

MEANS AND STANDARD DEVIATIONS ON THE WHOLE AND HALF-TESTS FOR NORTHERN WHITE SUBURBAN VERSUS NORTHERN BLACK URBAN

Grade 3.6

Test	Group	Whole Test		Half-Test	
		Mean*	SD	Mean*	SD
Vocabulary	I	34.8	7.3	19.7	4.3
	II	26.8	9.0	12.7	5.8
Comprehension	I	33.5	10.0	18.8	6.0
	II	23.5	9.9	11.1	6.7
Reading Total	I	68.5	16.0	36.0	9.6
	II	50.4	17.4	24.0	11.8
Computation	I	56.8	12.3	25.4	9.4
	II	52.5	15.0	21.4	10.8
Concepts & Problems	I	35.8	6.9	18.5	4.9
	II	27.2	7.6	13.4	5.5
Math Total	I	92.6	18.2	44.2	13.3
	II	80.5	20.0	33.4	15.5
Mechanics	I	32.3	15.0	16.4	9.2
	II	16.0	11.0	10.1	7.7
Usage & Structure	I	17.9	4.6	10.2	3.2
	II	13.6	4.7	7.2	3.6
Language Total	I	50.2	18.4	24.8	12.1
	II	30.2	14.0	16.5	10.0

*See footnote for Table 40.

Table 42

MEANS AND STANDARD DEVIATIONS ON THE WHOLE AND HALF-TESTS FOR NORTHERN WHITE SUBURBAN VERSUS NORTHERN BLACK URBAN

Grade 5.6

Test	Group	Whole Test		Half-Test	
		Mean*	SD	Mean*	SD
Vocabulary	I	30.5	6.8	14.1	4.8
	II	17.5	6.7	10.2	4.7
Comprehension	I	30.0	6.7	15.1	4.3
	II	18.7	6.4	11.3	4.7
Reading Total	I	60.6	12.7	28.9	8.6
	II	36.4	11.8	21.2	8.4
Computation	I	46.7	9.3	23.1	6.7
	II	34.9	11.0	21.1	7.9
Concepts & Problems	I	29.7	5.8	12.7	4.1
	II	19.9	6.6	11.3	4.6
Math Total	I	76.5	14.3	36.8	10.0
	II	55.1	16.1	33.0	11.2
Mechanics	I	50.1	14.4	24.6	9.1
	II	27.2	12.5	17.4	9.2
Usage & Structure	I	24.3	4.1	13.8	3.3
	II	19.1	4.5	11.1	3.4
Language Total	I	74.4	16.9	--	--
	II	46.7	15.1	26.5	11.9

*See footnote for Table 40.

Table 43

MEANS AND STANDARD DEVIATIONS ON THE WHOLE AND HALF-TESTS FOR NORTHERN WHITE SUBURBAN VERSUS NORTHERN BLACK URBAN

Grade 8.6

Test	Group	Whole Test		Half-Test	
		Mean*	SD	Mean*	SD
Vocabulary	I	28.3	6.3	15.3	4.0
	II	22.0	7.4	13.1	5.0
Comprehension	I	28.0	7.8	15.0	5.1
	II	20.5	6.6	11.5	4.7
Reading Total	I	56.3	13.2	30.5	8.5
	II	42.6	13.0	24.8	9.2
Computation	I	30.6	8.9	13.4	6.1
	II	24.3	8.9	11.0	6.1
Concepts & Problems	I	29.5	8.7	14.4	5.9
	II	21.7	7.8	12.8	5.5
Math Total	I	60.1	16.7	27.9	11.3
	II	46.2	15.6	23.8	10.9
Mechanics	I	45.5	11.1	20.4	7.2
	II	32.3	14.0	19.2	9.0
Usage & Structure	I	27.0	5.4	11.8	4.3
	II	23.1	5.7	12.0	4.1
Language Total	I	75.5	15.0	--	--
	II	55.6	17.4	--	--

*See footnote for Table 40.

Table 44

MEANS AND STANDARD DEVIATIONS ON THE WHOLE AND HALF-TESTS FOR SOUTHERN WHITE SUBURBAN VERSUS SOUTHERN BLACK RURAL

Grade 1.6

Test	Group	Whole Test		Half-Test	
		Mean*	SD	Mean*	SD
Vocabulary	III	67.5	11.1	38.2	7.3
	IV	56.1	20.0	25.5	13.0
Comprehension	III	8.1	3.6	5.1	2.7
	IV	9.0	6.6	4.0	3.8
Reading Total	III	75.8	12.9	44.9	9.0
	IV	65.3	25.2	29.0	16.1
Computation	III	28.0	8.5	14.4	5.6
	IV	17.1	12.0	8.0	6.8
Concepts & Problems	III	32.0	7.5	14.4	5.2
	IV	24.3	11.6	11.7	7.0
Math Total	III	60.1	14.2	30.0	9.5
	IV	41.6	22.8	19.7	13.7
Mechanics	III	18.0	7.7	9.7	5.5
	IV	12.2	11.1	6.5	6.8
Usage & Structure	III	14.3	3.1	8.0	2.2
	IV	8.0	5.5	3.7	3.3
Language Total	III	42.0	10.9	22.6	8.1
	IV	29.6	18.6	12.5	11.8

*See footnote for Table 40.

Table 45

MEANS AND STANDARD DEVIATIONS ON THE WHOLE AND HALF-TESTS FOR SOUTHERN WHITE SUBURBAN VERSUS SOUTHERN BLACK RURAL

Grade 3.6

Test	Group	Whole Test		Half-Test	
		Mean*	SD	Mean*	SD
Vocabulary	III	36.6	4.3	17.4	3.3
	IV	23.6	8.7	12.6	5.4
Comprehension	III	36.1	7.5	19.7	4.6
	IV	22.8	10.4	11.9	6.9
Reading Total	III	72.7	11.0	36.8	7.6
	IV	46.4	18.3	23.9	11.5
Computation	III	62.5	9.7	28.7	7.9
	IV	45.8	15.8	21.3	10.4
Concepts & Problems	III	38.1	5.1	19.0	3.8
	IV	26.8	10.3	14.1	6.3
Math Total	III	100.5	13.8	48.0	10.7
	IV	72.6	24.8	35.1	15.8
Mechanics	III	35.0	13.5	17.9	8.9
	IV	15.9	11.7	11.6	8.4
Usage & Structure	III	18.5	3.7	10.1	2.7
	IV	10.1	4.6	6.4	3.3
Language Total	III	53.5	16.0	26.0	11.4
	IV	26.0	15.3	16.5	10.9

*See footnote for Table 40.

Table 46

MEANS AND STANDARD DEVIATIONS ON THE WHOLE AND HALF-TESTS FOR SOUTHERN WHITE SUBURBAN VERSUS SOUTHERN BLACK RURAL

Grade 5.6

Test	Group	Whole Test Mean*	Whole Test SD	Half-Test Mean*	Half-Test SD
Vocabulary	III	31.1	6.5	14.9	4.3
	IV	15.1	7.6	8.4	5.0
Comprehension	III	29.8	7.1	16.7	4.3
	IV	16.0	7.1	10.0	5.3
Reading Total	III	60.9	12.8	32.0	7.8
	IV	31.2	13.7	18.5	9.5
Computation	III	44.5	10.4	21.1	7.2
	IV	27.7	11.4	16.1	8.3
Concepts & Problems	III	28.7	5.6	15.6	3.7
	IV	16.6	7.8	10.3	5.5
Math Total	III	73.2	15.0	35.3	10.4
	IV	44.3	18.4	26.3	13.0
Mechanics	III	48.1	14.2	24.5	8.9
	IV	24.4	13.8	14.4	9.7
Usage & Structure	III	24.7	4.6	15.0	3.3
	IV	19.5	4.5	10.0	3.4
Language Total	III	72.8	17.3	36.8	12.3
	IV	44.0	16.3	23.2	13.0

*See footnote for Table 40.

Table 47

MEANS AND STANDARD DEVIATIONS ON THE WHOLE AND HALF-TESTS FOR SOUTHERN WHITE SUBURBAN VERSUS SOUTHERN BLACK RURAL

Grade 8.6

Test	Group	Whole Test		Half-Test	
		Mean*	SD	Mean*	SD
Vocabulary	III	29.3	7.1	16.9	4.3
	IV	14.9	6.7	8.4	5.0
Comprehension	III	26.2	8.9	13.9	5.7
	IV	16.5	6.0	10.6	4.1
Reading Total	III	55.6	14.9	32.3	9.1
	IV	31.4	11.8	19.0	8.6
Computation	III	31.6	9.7	14.5	6.3
	IV	16.1	6.3	8.1	4.7
Concepts & Problems	III	30.0	9.6	16.0	6.0
	IV	17.2	7.3	10.1	5.1
Math Total	III	61.7	18.3	29.6	11.8
	IV	33.4	12.6	19.1	9.1
Mechanics	III	43.5	15.0	23.7	8.7
	IV	27.5	14.1	18.3	9.6
Usage & Structure	III	25.2	5.3	12.5	4.0
	IV	21.5	5.0	11.7	3.8
Language Total	III	68.7	18.8	37.7	13.7
	IV	48.9	17.3	26.0	13.3

*See footnote for Table 40.

Table 48

MEANS AND STANDARD DEVIATIONS ON THE WHOLE AND HALF-TESTS FOR SOUTHERN WHITE SUBURBAN VERSUS SOUTHERN BLACK RURAL

Grade 10.6

Test	Group	Whole Test Mean*	Whole Test SD	Half-Test Mean*	Half-Test SD
Vocabulary	III	27.1	7.5	12.8	4.9
	IV	15.3	5.5	9.2	4.2
Comprehension	III	28.0	7.7	15.5	5.0
	IV	18.8	6.3	10.7	4.9
Reading Total	III	55.1	14.5	27.5	9.2
	IV	34.1	11.0	20.4	8.6
Computation	III	34.0	8.9	15.7	5.9
	IV	21.3	11.0	9.0	7.1
Concepts & Problems	III	30.2	9.2	14.0	6.0
	IV	15.9	8.1	8.7	5.9
Math Total	III	64.2	17.1	29.3	11.3
	IV	37.3	18.4	17.4	12.8
Mechanics	III	58.6	12.8	30.0	8.4
	IV	36.7	17.5	22.0	11.6
Usage & Structure	III	30.2	6.2	14.4	5.1
	IV	23.5	4.7	11.0	4.0
Language Total	III	88.7	17.3	48.0	12.6
	IV	60.4	20.7	34.3	16.8

*See footnote for Table 40.

Table 49

MEANS AND STANDARD DEVIATIONS ON THE WHOLE AND HALF-TESTS FOR SOUTHERN BLACK RURAL VERSUS SOUTHERN WHITE RURAL

Grade 1.6

Test	Group	Whole Test		Half-Test	
		Mean*	SD	Mean*	SD
Vocabulary	IV	56.1	20.0	25.5	13.0
	V	55.1	17.2	31.0	11.7
Comprehension	IV	9.0	6.6	4.0	3.8
	V	7.5	4.7	3.6	3.0
Reading Total	IV	65.3	25.2	29.0	16.1
	V	62.7	20.3	35.3	13.6
Computation	IV	17.1	12.0	8.0	6.8
	V	19.8	11.5	11.0	6.8
Concepts & Problems	IV	24.3	11.6	11.7	7.0
	V	26.5	10.1	13.7	6.4
Math Total	IV	41.6	22.8	19.7	13.7
	V	46.7	20.5	24.9	12.9
Mechanics	IV	12.2	11.1	6.5	6.7
	V	11.0	6.1	7.0	4.4
Usage & Structure	IV	8.0	5.5	3.7	3.3
	V	8.7	4.3	5.5	2.9
Language Total	IV	29.6	18.6	12.5	11.8
	V	28.4	11.4	16.4	8.0

*See footnote for Table 40.

Table 50

MEANS AND STANDARD DEVIATIONS ON THE WHOLE AND HALF-TESTS FOR SOUTHERN BLACK RURAL VERSUS SOUTHERN WHITE RURAL

Grade 3.6

Test	Group	Whole Test		Half-Test	
		Mean*	SD	Mean*	SD
Vocabulary	IV	23.6	8.7	12.6	5.4
	V	30.6	8.1	15.9	4.9
Comprehension	IV	22.8	10.4	11.9	6.9
	V	30.2	10.0	16.7	6.3
Reading Total	IV	46.4	18.3	23.9	11.5
	V	60.8	17.0	32.1	10.8
Computation	IV	45.8	15.8	21.3	10.4
	V	54.1	13.3	25.2	9.1
Concepts & Problems	IV	26.8	10.3	14.1	6.3
	V	31.2	8.8	16.7	5.5
Math Total	IV	72.6	24.8	35.1	15.8
	V	85.5	20.7	42.1	13.6
Mechanics	IV	15.9	11.7	11.6	8.4
	V	26.4	13.6	14.6	8.9
Usage & Structure	IV	10.1	4.6	6.4	3.3
	V	14.2	4.7	8.5	3.3
Language Total	IV	26.0	15.3	16.5	10.9
	V	40.6	17.0	21.8	11.1

*See footnote for Table 40.

Table 51

MEANS AND STANDARD DEVIATIONS ON THE WHOLE AND HALF-TESTS FOR SOUTHERN BLACK RURAL VERSUS SOUTHERN WHITE RURAL

Grade 5.6

Test	Group	Whole Test Mean*	Whole Test SD	Half-Test Mean*	Half-Test SD
Vocabulary	IV	15.1	7.6	8.4	5.0
	V	22.1	8.7	12.6	5.1
Comprehension	IV	16.0	7.1	10.0	5.3
	V	22.2	7.5	12.7	4.9
Reading Total	IV	31.2	13.7	18.5	9.5
	V	44.3	15.4	25.1	9.5
Computation	IV	27.7	11.4	16.1	8.3
	V	40.6	11.7	23.2	8.6
Concepts & Problems	IV	16.6	7.8	10.3	5.5
	V	24.5	7.5	14.6	4.8
Math Total	IV	44.3	18.4	26.3	13.0
	V	65.2	18.2	38.2	12.6
Mechanics	IV	24.4	13.8	14.4	9.7
	V	42.4	16.3	26.3	10.2
Usage & Structure	IV	19.5	4.5	10.0	3.4
	V	21.7	4.2	13.0	3.6
Language Total	IV	44.0	16.3	23.2	13.0
	V	64.0	19.2	36.9	14.2

*See footnote for Table 40.

Table 52

MEANS AND STANDARD DEVIATIONS ON THE WHOLE AND HALF-TESTS FOR SOUTHERN BLACK RURAL VERSUS SOUTHERN WHITE RURAL

Grade 8.6

Test	Group	Whole Test		Half-Test	
		Mean*	SD	Mean*	SD
Vocabulary	IV	14.9	6.7	8.4	5.0
	V	24.0	8.0	13.6	5.3
Comprehension	IV	16.5	6.0	10.6	4.1
	V	23.5	8.1	13.0	5.2
Reading Total	IV	31.4	11.8	19.0	8.6
	V	47.5	15.1	27.4	9.8
Computation	IV	16.1	6.3	8.1	4.7
	V	27.7	9.8	13.0	6.3
Concepts & Problems	IV	17.2	7.3	10.1	5.1
	V	26.2	8.6	13.5	5.7
Math Total	IV	33.4	12.6	19.1	9.1
	V	54.0	17.5	26.7	11.6
Mechanics	IV	27.5	14.1	18.3	9.6
	V	40.0	13.5	20.7	8.4
Usage & Structure	IV	21.5	5.0	11.7	3.8
	V	25.5	5.6	12.2	4.2
Language Total	IV	48.9	17.3	26.0	13.3
	V	65.6	17.3	33.6	12.5

*See footnote for Table 40.

Table 53

MEANS AND STANDARD DEVIATIONS ON THE WHOLE AND HALF-TESTS FOR SOUTHERN BLACK RURAL VERSUS SOUTHERN WHITE RURAL

Grade 10.6

Test	Group	Whole Test Mean*	Whole Test SD	Half-Test Mean*	Half-Test SD
Vocabulary	IV	15.3	5.5	9.2	4.2
	V	21.1	6.7	10.7	4.6
Comprehension	IV	18.8	6.3	10.7	4.9
	V	23.3	7.0	13.6	4.9
Reading Total	IV	34.1	11.0	20.4	8.6
	V	44.4	12.8	24.9	8.6
Computation	IV	21.3	11.0	9.0	7.1
	V	32.6	9.9	14.5	6.7
Concepts & Problems	IV	15.9	8.1	8.7	5.9
	V	25.3	8.8	11.7	6.0
Math Total	IV	37.3	18.4	17.4	12.8
	V	57.9	17.6	27.1	12.1
Mechanics	IV	36.7	17.5	22.0	11.6
	V	53.4	14.8	27.3	9.6
Usage & Structure	IV	23.5	4.7	11.0	4.0
	V	27.6	6.2	14.3	4.9
Language Total	IV	60.4	20.7	34.3	16.8
	V	81.1	19.2	42.8	14.1

*See footnote for Table 40.

Table 54

MEANS AND STANDARD DEVIATIONS ON THE WHOLE AND HALF-TESTS FOR SOUTHWESTERN MEXICAN-AMERICAN VERSUS SOUTHWESTERN ANGLO-AMERICAN

Grade 1.6

Test	Group	Whole Test		Half-Test	
		Mean*	SD	Mean*	SD
Vocabulary	VI	42.8	12.4	23.4	9.4
	VII	75.5	11.7	32.7	9.4
Comprehension	VI	5.7	2.5	3.2	2.1
	VII	10.9	6.0	6.0	3.8
Reading Total	VI	48.8	13.0	29.0	10.7
	VII	86.4	16.7	35.4	13.5
Computation	VI	14.9	9.2	7.9	5.9
	VII	29.7	8.9	15.1	6.1
Concepts & Problems	VI	18.6	7.7	9.8	5.6
	VII	33.3	8.1	16.3	5.3
Math Total	VI	33.4	15.2	18.4	10.7
	VII	63.0	15.1	30.8	10.4
Mechanics	VI	7.0	3.9	4.7	3.4
	VII	19.1	10.0	10.7	6.5
Usage & Structure	VI	6.5	3.1	3.6	2.4
	VII	13.4	3.8	7.3	2.6
Language Total	VI	18.8	7.3	11.3	6.2
	VII	42.3	14.3	22.1	10.3

*See footnote for Table 40.

Table 55

MEANS AND STANDARD DEVIATIONS ON THE WHOLE AND HALF-TESTS FOR SOUTHWESTERN MEXICAN-AMERICAN VERSUS SOUTHWESTERN ANGLO-AMERICAN

Grade 3.6

Test	Group	Whole Test		Half-Test	
		Mean*	SD	Mean*	SD
Vocabulary	VI	21.7	8.4	10.6	5.6
	VII	36.7	4.9	18.8	3.1
Comprehension	VI	16.5	9.0	9.1	6.1
	VII	36.1	7.9	19.7	4.8
Reading Total	VI	38.4	15.9	19.2	11.0
	VII	72.7	11.5	37.6	7.4
Computation	VI	48.0	12.9	22.1	9.8
	VII	66.1	8.0	32.5	5.7
Concepts & Problems	VI	22.6	8.8	11.9	6.2
	VII	37.8	5.0	19.4	3.7
Math Total	VI	70.6	19.6	34.3	14.4
	VII	103.8	11.7	52.2	8.4
Mechanics	VI	18.8	10.6	13.4	8.1
	VII	34.2	13.0	18.3	8.5
Usage & Structure	VI	10.6	5.1	6.6	3.9
	VII	18.4	3.8	9.9	2.7
Language Total	VI	29.5	14.4	20.7	11.2
	VII	52.6	15.6	25.6	10.7

*See footnote for Table 40.

Table 56

MEANS AND STANDARD DEVIATIONS ON THE WHOLE AND HALF-TESTS FOR SOUTHWESTERN MEXICAN-AMERICAN VERSUS SOUTHWESTERN ANGLO-AMERICAN

Grade 5.6

Test	Group	Whole Test		Half-Test	
		Mean*	SD	Mean*	SD
Vocabulary	VI	15.8	7.1	9.4	4.9
	VII	30.7	6.8	13.9	4.7
Comprehension	VI	18.2	6.1	12.0	4.8
	VII	30.2	6.6	15.2	4.2
Reading Total	VI	34.0	12.2	21.6	9.0
	VII	60.9	12.6	29.3	8.2
Computation	VI	33.9	10.1	21.8	7.6
	VII	45.7	8.7	23.0	6.2
Concepts & Problems	VI	18.3	6.6	12.0	4.9
	VII	29.3	5.4	13.5	4.1
Math Total	VI	52.2	15.9	34.1	12.1
	VII	45.0	13.1	37.5	9.4
Mechanics	VI	27.3	11.9	18.6	9.2
	VII	46.5	14.2	22.0	9.2
Usage & Structure	VI	19.3	4.0	11.0	3.6
	VII	23.8	4.8	14.6	3.6
Language Total	VI	46.6	13.9	26.3	11.7
	VII	70.4	16.8	36.2	12.2

*See footnote for Table 40.

Table 57

MEANS AND STANDARD DEVIATIONS ON THE WHOLE AND HALF-TESTS FOR SOUTHWESTERN MEXICAN-AMERICAN VERSUS SOUTHWESTERN ANGLO-AMERICAN

Grade 8.6

Test	Group	Whole Test		Half-Test	
		Mean*	SD	Mean*	SD
Vocabulary	VI	19.9	7.9	11.4	5.4
	VII	31.9	5.8	15.6	4.1
Comprehension	VI	21.1	7.4	11.4	5.0
	VII	31.0	8.0	15.4	5.3
Reading Total	VI	41.0	14.6	23.6	10.1
	VII	62.9	13.1	30.9	8.8
Computation	VI	25.5	9.4	11.6	6.3
	VII	36.9	8.7	17.8	5.7
Concepts & Problems	VI	23.1	9.3	11.3	6.2
	VII	36.5	7.7	17.3	5.4
Math Total	VI	48.7	17.7	23.8	11.9
	VII	73.5	15.2	35.7	10.3
Mechanics	VI	40.1	13.0	22.1	7.8
	VII	52.7	12.6	25.6	8.4
Usage & Structure	VI	24.2	5.4	10.4	4.1
	VII	29.2	6.3	14.3	4.9
Language Total	VI	64.4	16.8	33.1	11.9
	VII	81.9	17.5	42.7	12.6

*See footnote for Table 40.

Table 58

RELIABILITY COEFFICIENTS AND STANDARDIZED RELIABILITY COEFFICIENTS ON THE WHOLE AND HALF-TESTS FOR NORTHERN WHITE SUBURBAN VERSUS NORTHERN BLACK URBAN

Grade 1.6

Test	Group	Whole Test		Half-Test	
		KR 20	100-Item KR 20	KR 20	100-Item KR 20
Vocabulary	I	.906	.913	.908	.914
	II	.931	.937	.925	.931
Comprehension	I	.749	.926	.764	.931
	II	.668	.893	.639	.884
Reading Total	I	.916	.904	.916	.904
	II	.930	.920	.933	.923
Computation	I	.938	.974	.939	.975
	II	.938	.974	.924	.968
Concepts & Problems	I	.866	.932	.840	.918
	II	.836	.916	.820	.906
Math Total	I	.939	.947	.941	.949
	II	.934	.942	.938	.946
Mechanics	I	.905	.962	.904	.962
	II	.911	.964	.910	.964
Usage & Structure	I	.760	.941	.784	.948
	II	.728	.931	.734	.933
Language Total	I	.911	.933	--	--
	II	.902	.926	--	--

Table 59

RELIABILITY COEFFICIENTS AND STANDARDIZED RELIABILITY COEFFICIENTS
ON THE WHOLE AND HALF-TESTS FOR
NORTHERN WHITE SUBURBAN VERSUS NORTHERN BLACK URBAN

Grade 3.6

Test	Group	Whole Test		Half-Test	
		KR 20	100-Item KR 20	KR 20	100-Item KR 20
Vocabulary	I	.944	.977	.945	.977
	II	.925	.969	.912	.963
Comprehension	I	.944	.974	.947	.975
	II	.917	.961	.919	.962
Reading Total	I	.965	.970	.960	.966
	II	.951	.958	.950	.957
Computation	I	.949	.982	.950	.963
	II	.959	.970	.958	.970
Concepts & Problems	I	.897	.951	.897	.951
	II	.873	.939	.869	.936
Math Total	I	.961	.954	.960	.953
	II	.958	.951	.961	.954
Mechanics	I	.957	.972	.945	.963
	II	.932	.954	.920	.946
Usage & Structure	I	.841	.955	.864	.962
	II	.804	.943	.828	.871
Language Total	I	.960	.964	.954	.958
	II	.934	.940	.931	.937

Table 60

RELIABILITY COEFFICIENTS AND STANDARDIZED RELIABILITY COEFFICIENTS
ON THE WHOLE AND HALF-TESTS FOR
NORTHERN WHITE SUBURBAN VERSUS NORTHERN BLACK URBAN

Grade 5.6

Test	Group	Whole Test		Half-Test	
		KR 20	100-Item KR 20	KR 20	100-Item KR 20
Vocabulary	I	.888	.952	.880	.948
	II	.837	.925	.842	.930
Comprehension	I	.866	.939	.831	.921
	II	.802	.906	.823	.918
Reading Total	I	.930	.942	.917	.931
	II	.888	.906	.893	.910
Computation	I	.895	.927	.884	.918
	II	.916	.941	.920	.945
Concepts & Problems	I	.847	.933	.811	.914
	II	.856	.937	.841	.929
Math Total	I	.930	.924	.919	.914
	II	.936	.931	.934	.929
Mechanics	I	.937	.949	.922	.937
	II	.914	.930	.917	.933
Usage & Structure	I	.556	.753	.675	.835
	II	.591	.779	.632	.826
Language Total	I	.927	.913	--	--
	II	.904	.886	.923	.908

Table 61

RELIABILITY COEFFICIENTS AND STANDARDIZED RELIABILITY COEFFICIENTS
ON THE WHOLE AND HALF-TESTS FOR
NORTHERN WHITE SUBURBAN VERSUS NORTHERN BLACK URBAN

Grade 8.6

Test	Group	Whole Test		Half-Test	
		KR 20	100-Item KR 20	KR 20	100-Item KR 20
Vocabulary	I	.863	.940	.841	.929
	II	.868	.942	.875	.946
Comprehension	I	.874	.939	.856	.929
	II	.805	.901	.798	.898
Reading Total	I	.924	.934	.912	.924
	II	.907	.920	.910	.922
Computation	I	.906	.953	.898	.948
	II	.902	.950	.896	.947
Concepts & Problems	I	.887	.940	.876	.934
	II	.850	.924	.847	.917
Math Total	I	.942	.943	.934	.935
	II	.930	.931	.928	.929
Mechanics	I	.907	.931	.891	.920
	II	.943	.958	.933	.951
Usage & Structure	I	.656	.792	.731	.844
	II	.716	.834	.724	.840
Language Total	I	.902	.883	--	--
	II	.930	.916	.930	--

Table 62

RELIABILITY COEFFICIENTS AND STANDARDIZED RELIABILITY COEFFICIENTS ON THE WHOLE AND HALF-TESTS FOR SOUTHERN WHITE SUBURBAN VERSUS SOUTHERN BLACK RURAL

Grade 1.6

Test	Group	Whole Test		Half-Test	
		KR 20	100-Item KR 20	KR 20	100-Item KR 20
Vocabulary	III	.913	.919	.915	.921
	IV	.965	.968	.960	.963
Comprehension	III	.650	.886	.686	.901
	IV	.911	.977	.890	.971
Reading Total	III	.907	.894	.913	.900
	IV	.971	.967	.966	.961
Computation	III	.920	.966	.926	.969
	IV	.960	.984	.947	.978
Concepts & Problems	III	.869	.934	.835	.915
	IV	.941	.971	.921	.961
Math Total	III	.932	.940	.924	.933
	IV	.973	.976	.966	.970
Mechanics	III	.897	.958	.896	.958
	IV	.961	.985	.958	.984
Usage & Structure	III	.702	.922	.759	.940
	IV	.900	.978	.876	.972
Language Total	III	.899	.924	.906	.930
	IV	.968	.976	.968	.976

Table 63

RELIABILITY COEFFICIENTS AND STANDARDIZED RELIABILITY COEFFICIENTS ON THE WHOLE AND HALF-TESTS FOR SOUTHERN WHITE SUBURBAN VERSUS SOUTHERN BLACK RURAL

Grade 3.6

Test	Group	Whole Test		Half-Test	
		KR 20	100-Item KR 20	KR 20	100-Item KR 20
Vocabulary	III	.875	.946	.858	.938
	IV	.904	.959	.895	.955
Comprehension	III	.906	.955	.909	.957
	IV	.922	.963	.924	.964
Reading Total	III	.935	.944	.936	.945
	IV	.953	.960	.947	.955
Computation	III	.936	.953	.940	.956
	IV	.956	.968	.952	.965
Concepts & Problems	III	.843	.923	.842	.922
	IV	.930	.967	.913	.959
Math Total	III	.946	.937	.946	.937
	IV	.969	.964	.964	.958
Mechanics	III	.947	.964	.941	.960
	IV	.942	.961	.933	.955
Usage & Structure	III	.768	.930	.780	.934
	IV	.786	.936	.797	.940
Language Total	III	.948	.952	.947	.952
	IV	.946	.951	.944	.949

Table 64

RELIABILITY COEFFICIENTS AND STANDARDIZED RELIABILITY COEFFICIENTS ON THE WHOLE AND HALF-TESTS FOR SOUTHERN WHITE SUBURBAN VERSUS SOUTHERN BLACK RURAL

Grade 5.6

Test	Group	Whole Test		Half-Test	
		KR 20	100-Item KR 20	KR 20	100-Item KR 20
Vocabulary	III	.882	.949	.860	.939
	IV	.875	.946	.863	.940
Comprehension	III	.881	.946	.865	.938
	IV	.840	.926	.863	.937
Reading Total	III	.932	.944	.918	.932
	IV	.918	.932	.919	.933
Computation	III	.911	.938	.897	.928
	IV	.917	.942	.923	.946
Concepts & Problems	III	.840	.929	.827	.923
	IV	.889	.952	.889	.952
Math Total	III	.934	.929	.924	.918
	IV	.948	.944	.948	.944
Mechanics	III	.936	.948	.920	.935
	IV	.934	.946	.933	.946
Usage & Structure	III	.658	.824	.708	.855
	IV	.587	.776	.666	.829
Language Total	III	.931	.918	.934	.921
	IV	.919	.904	.941	.929

Table 65

RELIABILITY COEFFICIENTS AND STANDARDIZED RELIABILITY COEFFICIENTS
ON THE WHOLE AND HALF-TESTS FOR
SOUTHERN WHITE SUBURBAN VERSUS SOUTHERN BLACK RURAL

Grade 8.6

Test	Group	Whole Test		Half-Test	
		KR 20	100-Item KR 20	KR 20	100-Item KR 20
Vocabulary	III	.901	.959	.906	.960
	IV	.836	.927	.856	.937
Comprehension	III	.902	.953	.882	.943
	IV	.776	.885	.753	.871
Reading Total	III	.941	.950	.935	.944
	IV	.888	.903	.892	.907
Computation	III	.920	.960	.904	.951
	IV	.842	.917	.836	.914
Concepts & Problems	III	.911	.953	.895	.945
	IV	.849	.918	.843	.915
Math Total	III	.953	.954	.943	.944
	IV	.909	.911	.907	.909
Mechanics	III	.952	.965	.939	.955
	IV	.947	.961	.941	.957
Usage & Structure	III	.659	.794	.715	.834
	IV	.625	.769	.694	.819
Language Total	III	.940	.928	.954	.944
	IV	.932	.918	.949	.938

Table 66

RELIABILITY COEFFICIENTS AND STANDARDIZED RELIABILITY COEFFICIENTS ON THE WHOLE AND HALF-TESTS FOR SOUTHERN WHITE SUBURBAN VERSUS SOUTHERN BLACK RURAL

Grade 10.6

Test	Group	Whole Test		Half-Test	
		KR 20	100-Item KR 20	KR 20	100-Item KR 20
Vocabulary	III	.894	.955	.871	.944
	IV	.754	.885	.785	.901
Comprehension	III	.861	.932	.846	.924
	IV	.766	.879	.808	.903
Reading Total	III	.931	.941	.919	.930
	IV	.859	.878	.883	.899
Computation	III	.906	.953	.886	.942
	IV	.939	.970	.933	.967
Concepts & Problems	III	.899	.947	.880	.936
	IV	.878	.935	.889	.941
Math Total	III	.944	.945	.934	.935
	IV	.953	.954	.955	.956
Mechanics	III	.937	.949	.928	.942
	IV	.960	.968	.957	.965
Usage & Structure	III	.741	.841	.807	.886
	IV	.524	.671	.675	.794
Language Total	III	.931	.910	.941	.922
	IV	.945	.928	.964	.952

Table 67

RELIABILITY COEFFICIENTS AND STANDARDIZED RELIABILITY COEFFICIENTS ON THE WHOLE AND HALF-TESTS FOR SOUTHERN BLACK RURAL VERSUS SOUTHERN WHITE RURAL

Grade 1.6

Test	Group	Whole Test		Half-Test	
		KR 20	100-Item KR 20	KR 20	100-Item KR 20
Vocabulary	IV	.965	.979	.960	.963
	V	.951	.955	.951	.955
Comprehension	IV	.911	.977	.890	.972
	V	.807	.946	.798	.943
Reading Total	IV	.971	.966	.966	.961
	V	.954	.548	.951	.943
Computation	IV	.960	.984	.947	.978
	V	.954	.981	.944	.977
Concepts & Problems	IV	.941	.971	.921	.961
	V	.920	.960	.901	.950
Math Total	IV	.973	.977	.966	.970
	V	.965	.970	.960	.965
Mechanics	IV	.961	.985	.958	.984
	V	.840	.932	.833	.930
Usage & Structure	IV	.900	.978	.876	.972
	V	.800	.952	.811	.955
Language Total	IV	.968	.977	.968	.977
	V	.902	.926	.894	.921

Table 68

RELIABILITY COEFFICIENTS AND STANDARDIZED RELIABILITY COEFFICIENTS ON THE WHOLE AND HALF-TESTS FOR SOUTHERN BLACK RURAL VERSUS SOUTHERN WHITE RURAL

Grade 3.6

Test	Group	Whole Test		Half-Test	
		KR 20	100-Item KR 20	KR 20	100-Item KR 20
Vocabulary	IV	.904	.960	.895	.955
	V	.926	.969	.916	.964
Comprehension	IV	.922	.963	.924	.965
	V	.929	.967	.930	.967
Reading Total	IV	.953	.960	.947	.954
	V	.957	.963	.954	.961
Computation	IV	.956	.968	.952	.965
	V	.952	.965	.944	.959
Concepts & Problems	IV	.930	.967	.913	.959
	V	.915	.960	.901	.952
Math Total	IV	.969	.964	.964	.958
	V	.964	.961	.958	.951
Mechanics	IV	.942	.964	.933	.955
	V	.945	.961	.934	.955
Usage & Structure	IV	.786	.936	.797	.940
	V	.812	.945	.828	.951
Language Total	IV	.946	.951	.944	.949
	V	.949	.954	.941	.946

Table 69

RELIABILITY COEFFICIENTS AND STANDARDIZED RELIABILITY COEFFICIENTS ON THE WHOLE AND HALF-TESTS FOR SOUTHERN BLACK RURAL VERSUS SOUTHERN WHITE RURAL

Grade 5.6

Test	Group	Whole Test		Half-Test	
		KR 20	100-Item KR 20	KR 20	100-Item KR 20
Vocabulary	IV	.875	.946	.863	.940
	V	.911	.962	.886	.951
Comprehension	IV	.840	.926	.863	.937
	V	.861	.937	.849	.931
Reading Total	IV	.918	.932	.919	.933
	V	.938	.948	.927	.939
Computation	IV	.917	.971	.923	.947
	V	.927	.949	.938	.957
Concepts & Problems	IV	.889	.953	.889	.953
	V	.892	.954	.889	.953
Math Total	IV	.948	.944	.948	.944
	V	.951	.947	.954	.950
Mechanics	IV	.934	.946	.933	.945
	V	.951	.961	.942	.953
Usage & Structure	IV	.587	.776	.666	.829
	V	.547	.746	.702	.852
Language Total	IV	.919	.904	.941	.929
	V	.941	--	.951	--

Table 70

RELIABILITY COEFFICIENTS AND STANDARDIZED RELIABILITY COEFFICIENTS ON THE WHOLE AND HALF-TESTS FOR SOUTHERN BLACK RURAL VERSUS SOUTHERN WHITE RURAL

Grade 8.6

Test	Group	Whole Test		Half-Test	
		KR 20	100-Item KR 20	KR 20	100-Item KR 20
Vocabulary	IV	.836	.927	.856	.937
	V	.898	.956	.898	.956
Comprehension	IV	.776	.885	.753	.872
	V	.876	.940	.852	.927
Reading Total	IV	.888	.903	.892	.907
	V	.935	.944	.928	.938
Computation	IV	.842	.917	.836	.914
	V	.917	.958	.898	.948
Concepts & Problems	IV	.849	.918	.843	.914
	V	.883	.937	.869	.929
Math Total	IV	.909	.911	.907	.909
	V	.945	.946	.937	.938
Mechanics	IV	.947	.961	.941	.957
	V	.938	.954	.925	.945
Usage & Structure	IV	.625	.770	.694	.819
	V	.689	.815	.728	.843
Language Total	IV	.932	.918	.949	.939
	V	.928	.913	.939	.927

Table 71

RELIABILITY COEFFICIENTS AND STANDARDIZED RELIABILITY COEFFICIENTS ON THE WHOLE AND HALF-TESTS FOR SOUTHERN BLACK RURAL VERSUS SOUTHERN WHITE RURAL

Grade 10.6

Test	Group	Whole Test		Half-Test	
		KR 20	100-Item KR 20	KR 20	100-Item KR 20
Vocabulary	IV	.754	.884	.785	.901
	V	.843	.930	.826	.922
Comprehension	IV	.766	.879	.808	.904
	V	.814	.906	.819	.910
Reading Total	IV	.859	.877	.883	.899
	V	.901	.915	.895	.910
Computation	IV	.939	.970	.933	.967
	V	.923	.961	.916	.958
Concepts & Problems	IV	.878	.935	.889	.941
	V	.891	.942	.881	.936
Math Total	IV	.953	.954	.955	.956
	V	.947	.948	.944	.945
Mechanics	IV	.961	.969	.957	.965
	V	.950	.960	.940	.951
Usage & Structure	IV	.524	.670	.675	.793
	V	.728	.832	.792	.875
Language Total	IV	.945	.927	.964	.952
	V	.940	.921	.951	.935

Table 72

RELIABILITY COEFFICIENTS AND STANDARDIZED RELIABILITY COEFFICIENTS ON THE WHOLE AND HALF-TESTS FOR SOUTHWESTERN MEXICAN-AMERICAN VERSUS SOUTHWESTERN ANGLO-AMERICAN

Grade 1.6

Test	Group	Whole Test		Half-Test	
		KR 20	100-Item KR 20	KR 20	100-Item KR 20
Vocabulary	VI	.895	.903	.909	.916
	VII	.935	.940	.934	.941
Comprehension	VI	.333	.675	.533	.826
	VII	.883	.969	.875	.967
Reading Total	VI	.880	.863	.910	.897
	VII	.952	.945	.953	.946
Computation	VI	.926	.969	.916	.965
	VII	.939	.975	.952	.980
Concepts & Problems	VI	.853	.925	.857	.927
	VII	.890	.945	.864	.931
Math Total	VI	.934	.942	.936	.944
	VII	.945	.952	.942	.949
Mechanics	VI	.675	.845	.767	.897
	VII	.946	.979	.941	.977
Usage & Structure	VI	.582	.874	.703	.922
	VII	.786	.948	.803	.953
Language Total	VI	.786	.834	.843	.880
	VII	.944	.958	.949	.962

Table 73

RELIABILITY COEFFICIENTS AND STANDARDIZED RELIABILITY COEFFICIENTS ON THE WHOLE AND HALF-TESTS FOR SOUTHWESTERN MEXICAN-AMERICAN VERSUS SOUTHWESTERN ANGLO-AMERICAN

Grade 3.6

Test	Group	Whole Test		Half-Test	
		KR 20	100-Item KR 20	KR 20	100-Item KR 20
Vocabulary	VI	.900	.957	.897	.956
	VII	.910	.962	.936	.973
Comprehension	VI	.896	.950	.895	.950
	VII	.916	.960	.921	.963
Reading Total	VI	.938	.947	.939	.948
	VII	.942	.950	.941	.949
Computation	VI	.938	.955	.944	.959
	VII	.933	.951	.932	.950
Concepts & Problems	VI	.898	.951	.900	.952
	VII	.831	.916	.830	.916
Math Total	VI	.951	.943	.954	.947
	VII	.936	.926	.936	.926
Mechanics	VI	.917	.944	.920	.946
	VII	.938	.958	.930	.953
Usage & Structure	VI	.829	.951	.864	.962
	VII	.779	.934	.780	.934
Language Total	VI	.932	.938	.938	.943
	VII	.942	.947	.938	.943

Table 74

RELIABILITY COEFFICIENTS AND STANDARDIZED RELIABILITY COEFFICIENTS ON THE WHOLE AND HALF-TESTS FOR SOUTHWESTERN MEXICAN-AMERICAN VERSUS SOUTHWESTERN ANGLO-AMERICAN

Grade 5.6

Test	Group	Whole Test		Half-Test	
		KR 20	100-Item KR 20	KR 20	100-Item KR 20
Vocabulary	VI	.859	.938	.856	.937
	VII	.892	.954	.869	.943
Comprehension	VI	.779	.894	.834	.923
	VII	.863	.937	.827	.919
Reading Total	VI	.896	.913	.909	.924
	VII	.929	.941	.911	.926
Computation	VI	.897	.928	.912	.938
	VII	.878	.914	.863	.903
Concepts & Problems	VI	.853	.936	.869	.943
	VII	.824	.921	.824	.921
Math Total	VI	.932	.927	.941	.937
	VII	.915	.909	.909	.902
Mechanics	VI	.903	.921	.913	.929
	VII	.935	.947	.921	.936
Usage & Structure	VI	.467	.681	.652	.820
	VII	.686	.842	.737	.872
Language Total	VI	.882	.861	.916	.900
	VII	.926	.912	.933	.920

Table 75

RELIABILITY COEFFICIENTS AND STANDARDIZED RELIABILITY COEFFICIENTS ON THE WHOLE AND HALF-TESTS FOR SOUTHWESTERN MEXICAN-AMERICAN VERSUS SOUTHWESTERN ANGLO-AMERICAN

Grade 8.6

Test	Group	Whole Test		Half-Test	
		KR 20	100-Item KR 20	KR 20	100-Item KR 20
Vocabulary	VI	.882	.949	.885	.951
	VII	.880	.948	.864	.941
Comprehension	VI	.848	.925	.829	.915
	VII	.893	.949	.869	.936
Reading Total	VI	.926	.936	.924	.935
	VII	.935	.944	.923	.934
Computation	VI	.909	.954	.900	.949
	VII	.915	.957	.902	.950
Concepts & Problems	VI	.899	.947	.889	.941
	VII	.883	.938	.868	.929
Math Total	VI	.946	.947	.941	.942
	VII	.942	.943	.934	.935
Mechanics	VI	.933	.951	.913	.936
	VII	.938	.955	.932	.950
Usage & Structure	VI	.670	.802	.709	.830
	VII	.751	.858	.802	.890
Language Total	VI	.923	.908	.934	.923
	VII	.935	.922	.943	.931